

Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 4, 2002, 19:02:26 ; Search time 81 Seconds
 (without alignments)
 26.321 Million cell updates/sec

Title: US-09-4622-625-2_COPY_145_160
 Sequence: 1 RGFRLSNYEVKGHRDV 16

Scoring table: BLOSUM62
 Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
 Listing first 45 summaries

Database : A_Geneseq_101002,*

- 1: /SIDS2/gcldata/geneseqp-embl/AA1980.DAT;*
- 2: /SIDS2/gcldata/geneseqp-embl/AA1981.DAT;*
- 3: /SIDS2/gcldata/geneseqp-embl/AA1983.DAT;*
- 4: /SIDS2/gcldata/geneseqp-embl/AA1983.DAT;*
- 5: /SIDS2/gcldata/geneseqp-embl/AA1984.DAT;*
- 6: /SIDS2/gcldata/geneseqp-embl/AA1985.DAT;*
- 7: /SIDS2/gcldata/geneseqp-embl/AA1986.DAT;*
- 8: /SIDS2/gcldata/geneseqp-embl/AA1987.DAT;*
- 9: /SIDS2/gcldata/geneseqp-embl/AA1988.DAT;*
- 10: /SIDS2/gcldata/geneseqp-embl/AA1989.DAT;*
- 11: /SIDS2/gcldata/geneseqp-embl/AA1990.DAT;*
- 12: /SIDS2/gcldata/geneseqp-embl/AA1991.DAT;*
- 13: /SIDS2/gcldata/geneseqp-embl/AA1992.DAT;*
- 14: /SIDS2/gcldata/geneseqp-embl/AA1993.DAT;*
- 15: /SIDS2/gcldata/geneseqp-embl/AA1994.DAT;*
- 16: /SIDS2/gcldata/geneseqp-embl/AA1995.DAT;*
- 17: /SIDS2/gcldata/geneseqp-embl/AA1996.DAT;*
- 18: /SIDS2/gcldata/geneseqp-embl/AA1997.DAT;*
- 19: /SIDS2/gcldata/geneseqp-embl/AA1998.DAT;*
- 20: /SIDS2/gcldata/geneseqp-embl/AA1999.DAT;*
- 21: /SIDS2/gcldata/geneseqp-embl/AA2000.DAT;*
- 22: /SIDS2/gcldata/geneseqp-embl/AA2001.DAT;*
- 23: /SIDS2/gcldata/geneseqp-embl/AA2002.DAT;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	86	100.0	181	18	AAM23723	Murine granulocyte
2	86	100.0	182	20	AAY00770	Mouse tag7 clone P
3	72	83.7	190	18	AAW3722	Bovine granulocyte
4	64	74.4	191	20	AAY00771	Human tag7 clone P
5	64	74.4	196	21	AAB4022	Human PRO1269 prot
6	64	74.4	196	21	AAB5583	Htr4g protein enc0
7	64	74.4	196	21	AAY66964	Chondrosarcoma pep
8	64	74.4	196	21	AAY99400	Human PRO1269 (UNQ
9	64	74.4	196	22	AAB6149	Protein of the inv
10	54	62.8	81	23	ABP08874	Human ORFX protein

OS Mus musculus.
 XX Key Peptide
 FT FT /label= "MGP-A antimicrobial peptide (Claim 3)"
 XX Location/Qualifiers
 170..181 /note= "MGP-A antimicrobial peptide (Claim 3)"

Drosophila melanogaster

Helminth aminopeptidase

Novel human diastase

Bacillus sp. KSM-K

Human colon cancer

Human polypeptide

Wound healing tissue

Human full length

Human polypeptide

Human full length

Drosophila melanogaster

Novel human diastase

Zea mays protein f

Zea mays protein f

Human secreted protein f

Murine peptidoglycan

Aminopeptidase O12

Aminopeptidase O12

C3 vegetable PEP-C

Human immune/haeme

Human nervous system

Human secreted protein f

Helicobacter pylori

Drosophila melanogaster

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Carassius auratus

Novel human diastase

Helicobacter pylori

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

Arabidopsis thaliana

Human diacylglycerol acyltransferase

Novel human diastase

Human secreted protein f

XX WPI; 1997-424753/39.
DR N-PSDB; AAT78510.

XX Antimicrobial (poly)peptide(s) from bovine and murine granulocytes - useful therapeutically, as preservatives for food, in water treatment and in agriculture

XX Claim 9; Fig 5; 56pp; English.

CC This protein comprises the precursor of a novel, claimed antimicrobial peptide A or MGP-A (see AAW23725). Its amino acid sequence was deduced from a cDNA clone (see AAT78510) obtained from murine bone marrow. MGP-A and the bovine homologue, BGP-A (see AAW23724), exhibit activity against Gram-positive and Gram-negative bacteria, fungi and viruses, specifically Staphylococcus aureus, Escherichia coli, Candida albicans, Salmonella typhimurium and C. neoforomans (claimed). They can be used in human or veterinary medicine (particularly to treat disorders associated with lipopolysaccharides, e.g. sepsis and endotoxaemia) or as preservatives in food products or in water supplies (claimed). They can also be applied to crops to reduce post-harvest spoilage or expressed in transgenic plants to increase their disease resistance. They have low immunogenicity. Carboxamidated analogues of MGP-A and BGP-A may also be used.

XX Sequence 181 AA:

Query Match	Score	Length	Best Local Similarity	Pred.	No.	Mismatches	Indels	Gaps
Qy 1 RGFRLRSNEYVKGHDRV 16	100.0%	181	AC	5.5e-07	0	0	0	0;
Db 144 RGFRLRSNEYVKGHDRV 159			XX					

XX Sequence 182 AA:

Query Match	Score	Length	Best Local Similarity	Pred.	No.	Mismatches	Indels	Gaps
Qy 1 RGFRLRSNEYVKGHDRV 16	100.0%	182	AC	5.5e-07	0	0	0	0;
Db 145 RGFRLRSNEYVKGHDRV 160			XX					

RESULT 3
AAW23722 standard; Protein; 190 AA.

ID	AAW23722	standard	Protein	190 AA.
XX	AAW23722			
AC	AAW23722;			
DT	18-FEB-1998	(first entry)		
XX	Bovine granulocyte peptide A precursor	(antimicrobial BGP-A).		
DE	Bovine granulocyte peptide A precursor			
XX	Antimicrobial peptide; antibiotic; antibacterial; antifungal; fungicide; antiprotozoal; protozoacide; antiviral; virucide; bovine granulocyte peptide A; BGP-A; preservative; sepsis; bovine granulocyte peptide A; cattle.			
KW				
OS	Bos taurus.			
XX				
FH	Key	Location/Qualifiers		
FT	Peptide	1..21		
FT	Peptide	/label= Sig_Peptide		
FT	Peptide	22..177		
FT	Peptide	/label= Propeptide		
FT	Peptide	178..190		
FT	Peptide	/label= Mat_Peptide		
FT	Peptide	/note= "BGP-A antimicrobial peptide (Claim 1)"		
XX				
PN	WO9729765-A1.			
XX				
PD	21-AUG-1997.			
XX				
PF	13-FEB-1997;	97WO-US02218.		
XX				
PR	16-FEB-1996;	96US-0011834.		
XX				
PA	(REBC) UNIV CALIFORNIA.			
XX				
PI	Selsted ME;			
XX				
DR	WPI; 1997-424753/39.			
DR	N-PSDB; AAT78509.			
XX				
PT	Antimicrobial (poly)peptide(s) from bovine and murine granulocytes - useful therapeuticly as preservatives for food, in water treatment and in agriculture			
PT				
XX				
PS	Claim 8; Fig 4; 56pp; English.			
XX				

This protein comprises the precursor (see AAW23722) of a novel, claimed antimicrobial peptide from bovine neutrophils, designated bovine granulocyte Peptide A or BGP-A (see AAW33724). Its amino acid sequence was deduced from a cDNA clone (see AAT78503) obtained from bovine bone marrow. BGP-A and the murine homologue, MGP-A (see AAW23725), exhibit activity against Gram-positive and Gram-negative bacteria, fungi and viruses, specifically *Staphylococcus aureus*, *Escherichia coli*, *Candida albicans*, *Salmonella typhimurium* and C. neoforans (claimed). They can be used in human or veterinary medicine (particularly to treat disorders associated with lipopolysaccharides, e.g. sepsis and endotoxaemia) or as preservatives in food products or in water supplies (claimed). They can also be applied to crops to reduce post-harvest spoilage or expressed in transgenic plants to increase their disease resistance. They have low immunogenicity. Carboxamidated analogues of BGP-A and MGP-A may also be used.

XX Sequence 190 AA;
 Query Match 83.7%; Score 72; DB 18; Length 190;
 Best Local Similarity 81.2%; Pred. No. 0.00016; Indels 0; Gaps 0;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RGFLRSNEYEVKGHRDV 16
 Db 152 RGYLTPEYEVKGHRDV 167

RESULT 4
 AAY00771 standard; Protein; 191 AA.
 XX AC AAY00771;
 XX DT 18-MAY-1999 (first entry)
 XX DE Human tag7 clone protein sequence.

XX Tag7; tumour growth inhibitor; mammalian tumour; carcinoma; sarcoma; melanoma; leukaemia; apoptosis inducer; human.
 KW Homo sapiens.
 XX PN WO9902686-A1.
 XX PD 21-JAN-1999.
 XX PR 10-JUL-1998; 98WO-EP04287.
 XX PR 11-JUL-1997; 97US-0893724.

PA (BOEH) BOEHRINGER INGELHEIM INT GMBH.
 PI Georgiev G, Kiselev S, Ostermann E, Prokhorchouk E;
 XX DR WPI; 1999-120887/10.
 DR N-PSDB; AAC21820.

XX New nucleic acid encoding tag7 - used to inhibit tumour growth and induce apoptosis, for treatment of carcinoma, sarcoma, melanoma and leukaemia

PS Claim 23: Page 126-127; 138pp; English.

XX This sequence is the human tag7 of the invention. Cells containing the tag7 DNA sequence are used to express recombinant tag7. Tag7 is used to produce and purify antibodies; to inhibit growth of mammalian tumours, especially for treating carcinoma (of liver, ovary, breast, cervix, lung, prostate, colon/rectum, bladder, testis, stomach, pancreas, mouth, head, and neck, squamous cell carcinoma or teratocarcinoma), sarcoma (Kaposi's, osteo- or fibro-sarcomas), melanoma or leukemia; and as a molecular weight marker. The tag7 polypeptide inhibits tumour growth and induces apoptosis. The tag7 coding sequences are also useful as probes for gene

CC mapping and detection of tag7 gene expression, and as primers. Antibodies CC against tag7 are used as reagents for detecting tag7; as an antagonist of CC tag7; for isolating tag7 and therapeutically to inhibit or delay tumour CC metastasis.
 XX SQ Sequence 191 AA;
 Query Match 74.4%; Score 64; DB 20; Length 191;
 Best Local Similarity 75.0%; Pred. No. 0.0038; Mismatches 2; Indels 0; Gaps 0;
 Matches 12; Conservative 2; Indels 0; Gaps 0;

Qy 1 RGFLRSNEYEVKGHRDV 16
 Db 153 QGALRNEYEVKGHRDV 168

RESULT 5
 AAB24022
 ID AAB24022 standard; Protein; 196 AA.
 XX AC AAB24022;
 XX DT 25-JAN-2001 (first entry)
 XX DE Human PRO1269 protein sequence SEQ ID NO:7.
 XX KW Human; tumour; diagnosis; neoplastic disease; proliferation; cancer;
 KW identification; tumorigenesis; anticancer; detection.
 XX OS Homo sapiens.
 XX PN WO200053750-A1.
 XX PD 14-SEP-2000.
 XX PF 02-DEC-1999; 99WO-US28551.
 XX PR 08-MAR-1999; 99WO-US05028.
 PR 01-SEP-1999; 99WO-US20111.
 PR 29-OCT-1999; 99US-0162506.
 PR 30-NOV-1999; 99WO-US2813.
 PR 01-DEC-1999; 99WO-US28634.
 XX PA (GETH) GENENTECH INC.

XX PI Botstein D, Goddard A, Gurney AL, Roy MA, Watanabe CK, Wood WI;
 XX DR WPI; 2000-594320/56.
 DR N-PSDB; AAC58104.

XX PT Antibodies specific for PRO polypeptides, used to diagnose and inhibit the growth of tumors in mammals, and to identify inhibitors of PRO polypeptide activity or expression -
 XX PS Claim 61; Fig 4; 226pp. English.

XX The present invention describes an antibody that binds to a human protein (I) selected from: PRO381; PRO1269; PRO1410; PRO1755; PRO1780; PRO3434; PRO1927; PRO3567; PRO1295; PRO1303; PRO344; PRO4354; PRO397; PRO4407; PRO1555; PRO1096; PRO2038; and PRO262. (I) has antitumour activity and can be used to diagnose tumours in mammals, by detecting complex formation when the antibody is contacted with test cells. Increased expression of genes encoding (I) can also be detected to diagnose tumours. Agents which inhibit the activity of (I), especially the antibodies, or an antisense oligonucleotide which hybridises to genes encoding (I), can be used to inhibit tumour growth, preferably by inducing cell death. Methods from the present invention can be used to identify compounds which inhibit the biological activity of (I). AAC58109 to AAC58102 represent PCR primers and hybridisation probes used in examples from the present invention for human PRO sequences. AAC58103 to AAC58122 and AAC24040 represent human PRO polynucleotide and protein sequences given in the exemplification of the present invention.

XX PT Polynucleotide encoding peptidoglycan recognition protein-like protein,
PT antibodies specific to it useful for preventing, treating conditions
e.g. endotoxic shock and auto-immune disorders and infections in mammal

XX PS Claim 11; Fig 3; 191pp; English.

CC Novel human peptidoglycan recognition protein-like proteins (PGRP-W)
CC expressed by keratinocytes (PGRP-K), wound healing tissues (PGRP-W)
CC or chondrosarcoma tissue (PGRP-C) are disclosed. The PGRP DNA and
CC condition in a mammal (claimed). PGRP is useful in ameliorating a medical
CC system in such areas as immune recognition, antigen presentation and
CC immune system activation. Antibodies or antagonists directed against
CC these proteins may be useful in reducing or eliminating disorders
CC associated with tumour necrosis factor (TNF) and TNF-like cytokines,
CC such as endotoxic shock and autoimmune disorders and for treating
CC infectious diseases including silicosis, sarcoidosis and idiopathic
CC pulmonary fibrosis.

XX Sequence 196 AA;

Query Match	Score	DB	Length
Best Local Similarity	64	21	196
Matches	74.4%	Pred No	0.0039;
	12; Conservative	2; Mismatches	2;
		Indels	0;
		Gaps	0;

QY 1 RGFLRSNVEVKGRDV 16

Db 158 QSALRSNVYVLKGHRDV 173

RESULT 8

ID AAY99400

XX AAY99400 standard; Protein; 196 AA.

AC AAY99400;

XX DT 08-AUG-2000 (first entry)

DE Human PRO1269 (UNQ639) amino acid sequence SEQ ID NO:216.

XX KW Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;
XX transmembrane; secretion; immunoadhesion; pharmaceutical; screening.

OS Homo sapiens.

XX PN WO200012708-A2.

XX PD 09-MAR-2000.

XX PF 01-SEP-1999; 99WO-US20111.

PR 01-SEP-1998; 98US-00998716.

PR 01-SEP-1998; 98US-0099749.

PR 01-SEP-1998; 98US-0099750.

PR 02-SEP-1998; 98US-0099803.

PR 02-SEP-1998; 98US-0099801.

PR 02-SEP-1998; 98US-0099821.

PR 02-SEP-1998; 98US-0099843.

PR 02-SEP-1998; 98US-0099536.

PR 02-SEP-1998; 98US-0099596.

PR 02-SEP-1998; 98US-0099752.

PR 02-SEP-1998; 98US-0099808.

PR 02-SEP-1998; 98US-0099802.

PR 02-SEP-1998; 98US-0099642.

PR 10-SEP-1998; 98US-0099741.

PR 10-SEP-1998; 98US-0099754.

PR 10-SEP-1998; 98US-0099763.

PR 10-SEP-1998; 98US-0099750.

PR 10-SEP-1998; 98US-0099598.

PR 10-SEP-1998; 98US-0099602.

PR 10-SEP-1998; 98US-0099812.

PR 10-SEP-1998; 98US-0099816.

PR 15-SEP-1998; 98US-0099816.

PR 15-SEP-1998; 98US-0100385.

PR 15-SEP-1998; 98US-0100388.

PR 15-SEP-1998; 98US-0100390.

PR 16-SEP-1998; 98US-0100584.

PR 16-SEP-1998; 98US-0100627.

PR 16-SEP-1998; 98US-0100661.

PR 16-SEP-1998; 98US-0100662.

PR 16-SEP-1998; 98US-0100664.

PR 17-SEP-1998; 98US-0100683.

PR 17-SEP-1998; 98US-0100684.

PR 17-SEP-1998; 98US-0100710.

PR 17-SEP-1998; 98US-0100711.

PR 17-SEP-1998; 98US-0100919.

PR 17-SEP-1998; 98US-0100930.

PR 18-SEP-1998; 98US-0100848.

PR 18-SEP-1998; 98US-0100849.

PR 18-SEP-1998; 98US-0101014.

PR 18-SEP-1998; 98US-0101068.

PR 18-SEP-1998; 98US-0101071.

PR 22-SEP-1998; 98US-0101279.

PR 23-SEP-1998; 98US-0101471.

PR 23-SEP-1998; 98US-0101472.

PR 23-SEP-1998; 98US-0101474.

PR 23-SEP-1998; 98US-0101475.

PR 23-SEP-1998; 98US-0101476.

PR 23-SEP-1998; 98US-0101477.

PR 23-SEP-1998; 98US-0101479.

PR 24-SEP-1998; 98US-0101738.

PR 24-SEP-1998; 98US-0101741.

PR 24-SEP-1998; 98US-0101743.

PR 24-SEP-1998; 98US-0101915.

PR 24-SEP-1998; 98US-0101916.

PR 29-SEP-1998; 98US-0102207.

PR 29-SEP-1998; 98US-0102240.

PR 29-SEP-1998; 98US-0102307.

PR 29-SEP-1998; 98US-0102331.

PR 30-SEP-1998; 98US-0102484.

PR 30-SEP-1998; 98US-0102487.

PR 30-SEP-1998; 98US-0102570.

PR 30-SEP-1998; 98US-0102571.

PR 01-OCT-1998; 98US-0102634.

PR 02-OCT-1998; 98US-0102687.

PR 06-OCT-1998; 98US-0102935.

PR 06-OCT-1998; 98US-010358.

PR 07-OCT-1998; 98US-010349.

PR 07-OCT-1998; 98US-0103314.

PR 07-OCT-1998; 98US-0103315.

PR 07-OCT-1998; 98US-0103328.

PR 07-OCT-1998; 98US-0103395.

PR 14-OCT-1998; 98US-0103711.

PR 07-OCT-1998; 98US-0103496.

PR 08-OCT-1998; 98US-0103633.

PR 08-OCT-1998; 98US-0103678.

PR 08-OCT-1998; 98US-0103679.

PR 22-OCT-1998; 98US-0103711.

PR 22-OCT-1998; 98US-0104257.

PR 20-OCT-1998; 98US-0104987.

PR 20-OCT-1998; 98US-0105000.

PR 20-OCT-1998; 98US-0105002.

PR 21-OCT-1998; 98US-0105104.

PR 27-OCT-1998; 98US-0105882.

PR 27-OCT-1998; 98US-0106062.

PR 28-OCT-1998; 98US-0106023.

PR 28-OCT-1998; 98US-0106029.

PR 28-OCT-1998; 98US-0106030.

PR 28-OCT-1998; 98US-0106032.

PR 28-OCT-1998; 98US-0106033.

PR 28-OCT-1998; 98US-0106178.

PR 29-OCT-1998; 98US-0106248.
 PR 29-OCT-1998; 98US-0106344.
 PR 29-OCT-1998; 98US-0106350.
 PR 30-OCT-1998; 98US-0106434.
 PR 03-NOV-1998; 98US-0106856.
 PR 03-NOV-1998; 98US-0106912.
 PR 03-NOV-1998; 98US-0106925.
 PR 03-NOV-1998; 98US-0106939.
 PR 03-NOV-1998; 98US-0106942.
 PR 03-NOV-1998; 98US-0106944.
 PR 10-NOV-1998; 98US-0107733.
 PR 17-NOV-1998; 98US-0108715.
 PR 17-NOV-1998; 98US-0108719.
 PR 17-NOV-1998; 98US-0108737.
 PR 17-NOV-1998; 98US-0108748.
 PR 17-NOV-1998; 98US-0108801.
 PR 17-NOV-1998; 98US-0108812.
 PR 17-NOV-1998; 98US-0108816.
 PR 17-NOV-1998; 98US-0108827.
 PR 17-NOV-1998; 98US-0108835.
 PR 18-NOV-1998; 98US-0108848.
 PR 18-NOV-1998; 98US-0108849.
 PR 18-NOV-1998; 98US-0108850.
 PR 18-NOV-1998; 98US-0108851.
 PR 18-NOV-1998; 98US-0108852.
 PR 18-NOV-1998; 98US-0108858.
 PR 18-NOV-1998; 98US-0108904.
 XX (GETH) GENENTECH INC.
 PI Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
 XX WPI; 2000-23787/20.
 DR N-PSD1; AAA37082.
 XX PT New mammalian DNA sequences encoding transmembrane, receptor or secreted PRO polypeptides useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interactions
 PR Claim 12; Fig 122; 773pp; English.
 XX CC AAA37022 to AAA37144 encode the new isolated human transmembrane, receptor or secreted PRO proteins given to AAY99462. The transmembrane and receptor PRO proteins can be used for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interactions. The polypeptides and nucleotide sequences encoding them have various industrial applications, including uses as pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent PCR primers and hybridisation probes used in the isolation of the PRO polypeptides from the present invention.
 XX SQ Sequence 196 AA;
 Query Match 74.4%; Score 64; DB 21; Length 196;
 Best Local Similarity 75.0%; Pred. No. 0.0039;
 Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 RESULT 10
 ABP08874 standard; Protein; 81 AA.
 ID ABP08874
 XX AC
 XX DT 24-JUN-2002 (first entry)
 DE Human ORFX protein sequence SEQ ID NO:17730.
 DE Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 DE hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 DE degenerative disorder; osteoarthritis; neurodegenerative disorder;
 DE cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 DE hypertension; hypothyroidism; cholesterol ester storage disease;
 DE immune deficiency; immune disorder; infectious disease;
 DE autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.

XX
OS Homo sapiens.
XX
PN WO200192523-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US10836.
XX
PR 30-MAY-2000; 2000US-206132P.
PR 29-AUG-2000; 2000US-228716P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach MD;
XX
DR WPI: 2002-106308/14.
DR N-PSDB; ABN24626.
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders -
PS Disclosure; SEQ ID 17730; 1037pp; English.
XX
CC The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN2752 encode the human ORFX
CC proteins given in ABP001.0 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC systemic cytokine damage.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
XX
SQ Sequence 81 AA;

Query Match	62.8%	Score 54;	DB 23;	Length 81;
Best Local Similarity	64.3%	Pred. No. 0.078;		
Matches	9;	Conservative	3;	Mismatches 2;
				Indels 0;
				Gaps 0;

XX
AC ABB64581 standard; Protein; 190 AA.
XX
AC ABB64581;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 20535.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 1-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABLO8684.

XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 20535; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176/ABL3051), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB5737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
XX
SQ Sequence 190 AA;
Query Match 57.0%; Score 49; DB 22; Length 190;
Best Local Similarity 57.1%; Pred. No. 1.5;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
DT 31-MAY-1994 (first entry)
Qy 1 RGFLRSNEYVGHR 14
||:||:|| :|||
Db 152 RGYLKDNYTLFGHR 165

RESULT 12
AAR51282
ID AAR51282 standard; Protein; 977 AA.
XX
AC AAR51282;
XX
DE Helminth aminopeptidase H11-1 variant encoded by the H11-1 gene.
XX
KW Helminth; aminopeptidase; H11-3; H11-2; H11-1; integral membrane;
KW alpha-amino acyl peptide hydrolase (microsomal); multigene family;
KW antigen; vaccine; parasite; human; virus; microbe.
XX
OS Haemonchus contortus.
XX
PN WO9233542-A.
XX
PD 25-NOV-1993.
XX
PF 07-MAY-1993; 93WO-GB00943.
XX
PR 08-MAY-1992; 92GB-009993.
XX

RESULT 11
ABB64581
ID ABB64581 standard; Protein; 190 AA.
XX
AC ABB64581;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 20535.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;

PA (AGRIC-) AGRIC & FOOD RES COUNCIL.
 XX
 PI Graham M, Knox DP, Dunn EA, Newton SE, Oliver JJ;
 PI Smith TS;
 XX
 DR WPI; 1993-386574/46.
 DR N-PSDB; AAQ52491.
 XX
 PT New DNA encoding amino-peptidase from Haemonchus contortus and derived peptide(s) - useful in protective vaccines, for new vectors, transformed cells and oligosaccharide, for incorporation in virus or microbe
 XX
 PS Claim 6; Page 78-80; 137pp; English.
 XX
 The sequences given in AAR51280-82 represent variants of the helminth aminopeptidase H11D encoded by the genes H11-3, -2 and -1 respectively. H11D is a protein doublet which shows homology to a family of integral membrane aminopeptidases. The differences between the coding sequences and the translated products, can be attributed to different mRNAs of the multigene family, and also to different variants of the H11D-encoding sequence being present at different stages of the life cycle, or in strains differing in geographical origin. Antigenic fragments of these aminopeptidases may be used in vaccines to stimulate immune response against helminth parasites in humans or other animals. The DNA sequences encoding these proteins may be incorporated into a virus or microbe and used in a similar manner.
 XX
 Sequence 977 AA;
 Query Match 54.7%; Score 47; DB 14; Length 977;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 RGFRLRSNYEVKGHRDV 16
 Db 630 RAFCRSNYDANGWRNI 645

RESULT 13
 ABG26884
 ID ABG26884 standard; Protein; 334 AA.
 XX
 AC ABG26884;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #26875.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PR 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS91071.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess

PT biodiversity
 XX
 PS Claim 20; SEQ ID NO 57243; 103pp; English.
 XX
 The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridization probes, polymerase chain reaction (PCR), primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity amino acid sequences. ABG001010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 Sequence 334 AA;
 Query Match 53.5%; Score 46; DB 22; Length 334;
 Best Local Similarity 58.3%; Pred. No. 9.7;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 Qy 4 IRSNVEVKGHRD 15
 Db 292 LRDNFDFLVGHRD 303
 RESULT 14
 ARB73613
 ID ARB73613 standard; Protein; 432 AA.
 XX
 AC ARB73613;
 XX
 DT 10-AUG-2001 (first entry)
 XX
 DE Bacillus sp. KSM-KP43 origin of replication-associated protein.
 XX
 KW Origin of replication-associated protein; ori; plasmid construction; expression vector; recombinant expression.
 XX
 OS Bacillus sp. KSM-KP43.
 XX
 PN JP2001069986-A.
 XX
 PD 21-MAR-2001.
 PP 03-SEP-1999; 99JP-0250690.
 XX
 PR 03-SEP-1999; 99JP-0250690.
 XX
 PA (RAOS) KAO CORP.
 XX
 DR WPI; 2001-297392/31.
 DR N-PSDB; AAH33948.
 XX
 PT A Plasmid vector used for the preparation of a useful protein or a useful peptide -
 XX
 PS Claim 2; Page 6-7; 9pp; Japanese.
 XX
 CC The invention relates to a plasmid vector containing a 2.2 kb genomic fragment from *Bacillus* sp. KSM-KP43 (AAH23948) which contains the origin of replication (ori). This vector can be used for the

Copyright (c) 1993 - 2002 Compugen Ltd.

OM Protein - protein search, using sw model

Run on: November 4, 2002, 20:33:26 ; Search time 199 Seconds

Title: US-09-462-625-2_COPY_145_160
Perfect score: 86
Sequence: 1 RGFLRSNYEVKGHRDV 16

Scoring table: BLOSUM62

Gapext 0.5

Searched: 90412 seqs, 13869272 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaas/US08_NEW_PUB.pep;*
2: /cgn2_6/ptodata/1/pubpaas/PCT_NEW_PUB.pep;*
3: /cgn2_6/ptodata/1/pubpaas/US06_NEW_PUB.pep;*
4: /cgn2_6/ptodata/1/pubpaas/US05_NEW_PUB.pep;*
5: /cgn2_6/ptodata/1/pubpaas/US04_NEW_PUB.pep;*
6: /cgn2_6/ptodata/1/pubpaas/US07_NEW_PUB.pep;*
7: /cgn2_6/ptodata/1/pubpaas/PCRS_PUBCOMB.pep;*
8: /cgn2_6/ptodata/1/pubpaas/PUBCOMB.pep;*
9: /cgn2_6/ptodata/1/pubpaas/US08_NEW_PUB.pep;*
10: /cgn2_6/ptodata/1/pubpaas/US09_NEW_PUB.pep;*
11: /cgn2_6/ptodata/1/pubpaas/US05_PUBCOMB.pep;*
12: /cgn2_6/ptodata/1/pubpaas/US10_NEW_PUB.pep;*
13: /cgn2_6/ptodata/1/pubpaas/US06_NEW_PUB.pep;*
14: /cgn2_6/ptodata/1/pubpaas/US06_PUBCOMB.pep;*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	47.7	131	10 US-09-71-161A-148	Sequence 148, App
2	41	47.7	736	10 US-09-815-242-11293	Sequence 11293, A
3	41	47.7	736	10 US-09-815-242-111473	Sequence 11473, A
4	41	47.7	757	10 US-09-725-735A-20	Sequence 20, Appli
5	41	47.7	1065	10 US-09-711-161A-239	Sequence 239, App
6	37	43.0	101	10 US-09-724-877-1470	Sequence 1470, Ap
7	37	43.0	358	10 US-09-864-761-48087	Sequence 48087, A
8	36	41.9	26	10 US-09-854-761-3850	Sequence 38500, A
9	36	41.9	407	10 US-09-815-242-5348	Sequence 5348, Ap
10	36	41.9	444	10 US-09-815-242-12376	Sequence 12376, A
11	36	41.9	732	10 US-09-994-483-8	Sequence 8, Appli
12	36	41.9	777	10 US-09-925-301-1115	Sequence 1115, Ap
13	35	40.7	62	10 US-09-867-761-38482	Sequence 38482, A
14	35	40.7	457	10 US-09-898-70-24	Sequence 24, Appli
15	35	40.7	558	10 US-09-815-242-5665	Sequence 5665, Ap
16	35	40.7	605	10 US-09-815-242-1258	Sequence 12658, A
17	35	40.7	1457	10 US-09-77-316-2	Sequence 2, Appli
18	35	40.7	1460	10 US-09-315-242-11668	Sequence 13668, A
19	35	10	35	US-09-864-761-38250	Sequence 38250, A

ALIGNMENTS

```
RESULT 1
US-09-71-161A-148
; Sequence 148, Application US/09771161A
; Patent No. US/02011081A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09-771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIORITY APPLICATION NUMBER: 09/724,676
; PRIORITY FILING DATE: 2000-11-28
; PRIORITY APPLICATION NUMBER: 136776
; PRIORITY FILING DATE: 2000-06-15
; PRIORITY APPLICATION NUMBER: 135619
; PRIORITY FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 148
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-71-161A-148

Query Match          47.7%
Best Local Similarity 58.3%
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy      5 RSNYEVKGHRDV 16
Db      116 RQNKVIGHEDL 127

RESULT 2
US-09-815-242-11293
; Sequence 11293, Application US/09815242
; Patent No. US/2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
```

APPLICANT: Trawick, John D.
 APPLICANT: Carr, Grant J.
 APPLICANT: Yamamoto, Robert T.
 APPLICANT: Xu, H. Howard
 TITLE OF INVENTION: Identification of Essential Genes in
 TITLE OF INVENTION: Prokaryotes
 FILE REFERENCE: ELITRA.011A
 CURRENT APPLICATION NUMBER: US/09/815,242
 CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 14110
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 11293
 LENGTH: 736
 TYPE: PRT
 ORGANISM: Helicobacter pylori
 US-09-815-242-11293

Query Match 47.7%; Score 41; DB 10; Length 736;
 Best Local Similarity 55.0%; Pred. No. 20;
 Matches 11; Conservative 2; Mismatches 3; Indels 4; Gaps 2;

Qy 1 RGFLRSNEY--KGH-RDV 16
 : ||| : ||| :
 Db 17 KNFLDKNTVEVASKGHVRDL 36

RESULT 3
 US-09-815-242-11473
 Sequence 11473, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hasebeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 TITLE OF INVENTION: Identification of Essential Genes in
 TITLE OF INVENTION: prokaryotes
 FILE REFERENCE: ELITRA.011A
 CURRENT APPLICATION NUMBER: US/09/815,242
 CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-33
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 11473
 ; LENGTH: 736
 ; TYPE: PRT
 ; ORGANISM: Helicobacter pylori
 ; US-09-815-242-11473

Query Match 47.7%; Score 41; DB 10; Length 736;
 Best Local Similarity 55.0%; Pred. No. 20;
 Matches 11; Conservative 2; Mismatches 3; Indels 4; Gaps 2;

Qy 1 RGFLRSNEY--KGH-RDV 16
 : ||| : ||| :
 Db 17 KNFLDKNTVEVASKGHVRDL 36

RESULT 4
 US-09-725-735A-20
 Sequence 20, Application US/09725735A
 ; Patent No. US20020031797A1
 ; GENERAL INFORMATION:
 ; APPLICANT: RUSING, MATTHIAS
 ; APPLICANT: SCHWEINS, THOMAS
 ; APPLICANT: DREESLER, PETRA
 ; APPLICANT: STOCK, WOLFGANG
 ; APPLICANT: KIVI, THOMAS
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID ISOLATED FROM TETRAHYMENA WHICH
 ; TITLE OF INVENTION: CODES FOR A TRITERPENOID CYCLASE, ITS PRODUCTION AND
 ; TITLE OF INVENTION: USE
 ; FILE REFERENCE: 25426 0001
 ; CURRENT APPLICATION NUMBER: US/09/725,735A
 ; CURRENT FILING DATE: 2000-11-30
 ; PRIOR APPLICATION NUMBER: DE 199-57 889.3
 ; PRIOR FILING DATE: 1999-12-01
 ; NUMBER OF SEQ ID NOS: 23
 ; SEQ ID NO 20
 ; LENGTH: 757
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-09-725-735A-20

Query Match 47.7%; Score 41; DB 10; Length 757;
 Best Local Similarity 50.0%; Pred. No. 21;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RGFLRSNEYEVGHARDV 16
 : ||| : ||| :
 Db 48 RGFLDNFRVKGCSDL 63

RESULT 5
 US-09-71-161A-239
 Sequence 239, Application US/09771161A
 ; Patent No. US20020110811A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LEVINE, et al.
 ; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
 ; FILE REFERENCE: 802620-205.1
 ; CURRENT APPLICATION NUMBER: US/09/771,161A
 ; CURRENT FILING DATE: 2001-01-26
 ; PRIOR APPLICATION NUMBER: 09/724,676
 ; PRIOR FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: 136776
 ; PRIOR FILING DATE: 2000-06-15
 ; PRIOR APPLICATION NUMBER: 135619
 ; PRIOR FILING DATE: 2000-04-12
 ; NUMBER OF SEQ ID NOS: 273
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 239
 ; LENGTH: 1065
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens

US-09-771-161A-239

Query Match 47.7%; Score 41; DB 10; Length 1065;
 Best Local Similarity 58.3%; Pred. No. 30; Length 1065;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 5 RSNYEVKGHRDV 16
 Db 1050 RQNYKVIGHDL 1061

RESULT 6

US-09-764-877-1470

; Sequence 1470, Application US/09764877

; Patent No. US20020147140A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC005

; CURRENT APPLICATION NUMBER: US/09/7764, 877

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 4031

; SOFTWARE: PatentIn Ver. 2.0

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; SEQ ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

; ID NO: 1470

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 SEQ ID NO: 38500

LENGTH: 26

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AC005288.1

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 28

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 23

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 23

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 25

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 25

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 23

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 31

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 24

OTHER INFORMATION: EXPRESSED IN HUMAN HIT: BF433100.1, SIGNAL = 24.00e-03

US-09-864-761-38500

Query Match Score 36; DB 10; Length 26;
 Best Local Similarity 66.7%; Pred. No. 3.8;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RGFLRSNYEVKG 12

Db 15 RGILRNYVVCG 26

RESULT 9

US-09-815-242-5348

Sequence 5348, Application US/09815242

; Patent No. US2002061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.01A

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO: 12376

; LENGTH: 444

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

RESULT 11
 US-09-94-485-8
 ; Sequence 8, Application US/09994485

; Patent No. US200201429A1
; GENERAL INFORMATION:
; APPLICANT: Ryazanov, Alexey G.
; Hait, William N.
; Pavur, Karen S.
; TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.3.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/994,485
; FILING DATE: 27 Nov. US200201429A1-2001
; CLASSIFICATION: <Unknown>
; PRIORITY DATA:
; APPLICATION NUMBER: US/08/914,999
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 732 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Dictyostelium discoideum
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
; US-09-994-485-8

; Query Match 41.9%; Score 36; DB 10; Length 732;
; Best Local Similarity 66.7%; Pred. No. 1.4e+02;
; Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
; Qy 5 RSNEYKGH 13
; Db 613 RCNYTLKGH 621

; RESULT 12
; US-09-925-301-1115
; Sequence 1115, Application US/0923301
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PAL06
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US01/005882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1115
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (14)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (21)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (32)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-301-1115

; Query Match 41.9%; Score 36; DB 10; Length 777;
; Best Local Similarity 42.9%; Pred. No. 1.5e+02;
; Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
; Qy 3 FLRSNEYKGRHDV 16
; Db 413 FIRSEYRKGFQEV 426

; RESULT 13
; US-09-864-761-38482
; Sequence 38482, Application US/09864761
; Patent No. US20020148763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.-6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661

Gencore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run On: November 4, 2002, 20:31:22 ; Search time 50 Seconds
(without alignments)
30.763 Million cell updates/sec

Title: US-09-462-625-2_COPY_145_160
Perfect score: 86
Sequence: 1 RGFLRSNEYVKGRHDV 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 28324 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 28324

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR73:*

1: pir1;*
2: pir2;*
3: pir3;*
4: pir4;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8 Result No. Score Query Match Length DB ID Description

Result No.	Score	Query Match	Length	DB	ID	Description
1	45	52.3	471	2	SS1333	IgA-specific metalloendopeptidase homolog PH024
2	44.5	51.7	504	1	G71248	phosphoenolpyruvat phosphoenolpyruvat phosphoenolpyruvat phosphoenolpyruvat endopeptidase La (phosphoenolpyruvat phosphoenolpyruvat phosphoenolpyruvat phosphoenolpyruvat DNA topoisomerase DNA topoisomerase hypothetical prote hypothetical prote hypothetical prote probable phosphoenzyme endopeptidase La (ferric-pseudobacti endopeptidase La (endopeptidase La (Lon ATP-dependent proteinase, Lon fa phosphoenolpyruvat phosphoenolpyruvat thyrotropin-releas
3	43	50.0	960	2	S18240	
4	43	50.0	965	2	T09846	
5	43	50.0	967	2	JH0667	
6	42	48.8	787	2	A72230	
7	42	48.8	960	2	JH0381	
8	42	48.8	960	2	S31159	
9	42	48.8	964	1	QINT	
10	42	48.8	966	2	S18318	
11	42	48.8	967	2	S52853	
12	42	48.8	967	2	S25081	
13	41	47.7	736	2	B71972	
14	41	47.7	736	2	D64534	
15	41	47.7	830	2	T01058	
16	40.5	47.1	318	2	G75279	
17	40.5	47.1	469	2	B70201	
18	40.5	47.1	652	2	T19783	
19	40	46.5	328	2	T06547	
20	40	46.5	46.5	2	F83549	
21	40	46.5	806	1	E70131	
22	40	46.5	809	2	S32899	
23	40	46.5	819	1	C71527	
24	40	46.5	819	1	B72128	
25	40	46.5	819	2	E86494	
26	40	46.5	819	2	E81681	
27	40	46.5	966	1	QIXI1	
28	40	46.5	966	2	S26399	
29	40	46.5	1025	2	I59331	

ALIGNMENTS

RESULT 1

S61333

IgA-specific metalloendopeptidase (EC 3.4.24.13) homolog - Haemophilus influenzae (st

N; Alternative names: IgA1 protease

C; Species: Haemophilus influenzae

A; Variety: HK284

C; Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 08-Dec-2000

R; Lomholt, H.; Poulsen, K.; Mogens, K.

Mol. Microbiol. 15, 495-506, 1995

A; Title: Comparative characterization of the iga gene encoding IgA1 protease in Neiss

A; Accession: S61333

A; Status: preliminary; nucleic acid sequence not shown

A; Molecule type: DNA

A; Residues: 1-471 <LON>

A; Cross-references: EMBL:X82487; NID:9733205; PMID:7783620

A; Experimental source: strain HK284

C; Genetics:

A; Gene: iga

C; Superfamily: IgA-specific metalloendopeptidase

C; Keywords: hydrolase; metalloproteins

Query Match 52.3%; Score 45; DB 2; Length 471;

Best Local Similarity 61.5%; Pred. No. 9.8; Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 3 FLRSNYVKGRHD 15

Db 357 FFKGNYVKRGKTD 369

RESULT 2

G71248

tLDD homolog PH0246 - Pyrococcus horikoshii

C; Species: Pyrococcus horikoshii

C; Accession: G71248

R; Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Yamamoto, S., Se

M.; Ohfuki, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu

DNA Res. 5, 55-76, 1998

A; Title: Complete sequence and gene organization of the genome of a hyper-thermophilic

A; Reference number: A71000; MUID:98344137; PMID:9679194

A; Accession: G71248

A; Status: nucleic acid sequence not shown; translation not shown

A; Molecule type: DNA

A; Residues: 1-304 <KAW>

A; Cross-references: GB:Af00001; NID:93236128; PMID:93256635

A; Experimental source: strain OT3

A; Note: this accession replaces an interim accession for a sequence replaced by GenBa

C; Genetics:

A; Gene: PH0246

C; Superfamily: Escherichia coli tldD protein	RESULT 5
Query Match Score 44.5; DB 1; Length 504;	JH0667
Best Local Similarity 62.5%; Pred. No. 13;	phosphoenolpyruvate carboxylase (EC 4.1.1.31) C3-form - maize
Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;	C; Species: Zea mays (maize)
Qy 2 GFLRSNYEVKGH-RDV 16	C; Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 22-Jun-1999
Db 436 GYLVENGEIKGHLRDV 451	C; Accession: JH0667
	R; Kawamura, T.; Shigesada, K.; Toh, H.; Okumura, S.; Yanagisawa, S.; Izui, K.
	J. Biochem. 112, 147-154, 1992
	A; Title: Molecular evolution of phosphoenolpyruvate carboxylase for C4 photosynthesis
	A; Function:
	A; Reference number: JH0667; MUID: 93054411; PMID: 1429504
	A; Accession: JH0667
	A; Molecule type: mRNA
	A; Residues: 1-967 <KRW>
	A; Cross-references: GB:X61489; NID:9429148; PIDN:CAA43709.1; PMID:9429149
	A; Experimental source: root
	C; Comment: This enzyme catalyzes the carboxylation of phosphoenolpyruvate to form oxa
	C; Superfamily: phosphoenolpyruvate carboxylase
	C; Keywords: carbon-carbon lyase; carboxy-lyase
RESULT 3	
S18240	Query Match Score 43; DB 2; Length 967;
phosphoenolpyruvate carboxylase (EC 4.1.1.31) - sorghum	Best Local Similarity 47.6%; Pred. No. 46;
C; Species: Sorghum bicolor (sorghum)	Matches 10; Conservative 2; Mismatches 1; Indels 8; Gaps 1;
C; Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 22-Jun-1999	
Qy R; Lepiniec, L.; Santini, S.; Keryer, E.; Amitet, V.; Vidal, J.; Gadal, P.; Cretin, C.	
Plant Mol. Biol. 17, 1077-1079, 1991	
A; Title: Complete nucleotide sequence of one member of the Sorghum phosphoenolpyruvate c	
A; Reference Number: S18240; PMID: 1840886	
A; Molecule type: DNA	
A; Residues: 1-960 <LEP>	
A; Cross-references: ENB:X59925; NID:922532; PIDN:CAA42549.1; PMID:g22593	
A; Note: the authors translated the codon TRG for residue 395 as Phe, AAT for residue 696	
C; Genets: A; Introns: 51/3; 182/3; 210/3; 284/3; 319/3; 348/3; 400/3; 733/3; 862/3	
C; Superfamily: phosphoenolpyruvate carboxylase	
C; Keywords: carbon dioxide fixation; carbon-carbon lyase; carboxy-lyase	
Query Match Score 43; DB 2; Length 960;	RESULT 6
Best Local Similarity 47.6%; Pred. No. 46;	JH2230
Matches 10; Conservative 2; Mismatches 1; Indels 8; Gaps 1;	endopeptidase La (EC 3.4.21.53) - Thermotoga maritima (strain MSB8)
Qy 4 LRSNYE-----VKHRDV 16	N; Contains: adenosinetriphosphatase (EC 3.6.1.3)
Db 849 LRANYETEQKLQLQVAGHRDL 876	C; Species: Thermotoga maritima
	C; Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
	C; Accession: AT72230
	R; Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Dodson, M.L.; Haft, D.H.; Hic
	Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
	C.M.
	Nature 399, 323-329, 1999
	A; Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
	A; Reference number: AT72200; MUID: 99287316; PMID: 10360571
	A; Accession: AT7230
	A; Status: preliminary
	A; Molecule type: DNA
	A; Residues: 1-787 <ARN>
	A; Cross-references: GB:AE001806; GB:AE000512; NID:94982196; PIDN:AAD36700.1; PMID:9498
	A; Experimental source: strain MSB8
	C; Genetics:
	A; Gene: TM1633
	C; Superfamily: ATP-dependent serine proteinase
	C; Keywords: ATP; hydrolase; serine proteinase
RESULT 4	
T09846	Query Match Score 42; DB 2; Length 787;
phosphoenolpyruvate carboxylase (EC 4.1.1.31) 1 - upland cotton	Best Local Similarity 61.5%; Pred. No. 55;
C; Species: Gossypium hirsutum (upland cotton)	Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
C; Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 24-Nov-1999	
C; Accession: T09846	
A; Vojdani, F.; Wilkins, T. A.	
Submitted to the EMBL Data Library, June 1997	
A; Reference number: 216883	
A; Accession: T09846	
A; Status: preliminary; translated from GB/EMBL/DDBJ	
A; Molecule type: mRNA	
A; Residues: 1-965 <VO>>	
A; Cross-references: ENB:AE008939; NID:92266946; PMID:92266947	
C; Function: <PEPC>	
A; Description: catalyzes incorporation of a carbon dioxide molecule into phosphoenolpyru	
C; Pathway: tricarboxylic acid cycle	
A; Function: <CDF>	
A; Description: catalyzes the carboxylation (by carbon dioxide) of phosphoenolpyruvate to	
C; Pathway: carbon dioxide fixation	
C; Superfamily: phosphoenolpyruvate carboxylase	
C; Keywords: carbon dioxide fixation; carbon-carbon lyase; carboxy-lyase; phosphoprotein	
Query Match Score 43; DB 2; Length 965;	RESULT 7
Best Local Similarity 47.6%; Pred. No. 46;	JH0381
Matches 10; Conservative 2; Mismatches 1; Indels 8; Gaps 1;	phosphoenolpyruvate carboxylase (EC 4.1.1.31) - sorghum
Qy 4 LRSNYE-----VKHRDV 16	C; Species: Sorghum bicolor
Db 854 LRANYETOKRLQLQVAGHRDL 874	C; Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 18-Jul-2001
	C; Accession: JH0381; S16455
	R; Cretin, C.; Santini, S.; Keryer, E.; Lepiniec, L.; Tagu, D.; Vidal, J.; Gadal, P.
	Gene 99, 87-94, 1991
	A; Title: The phosphoenolpyruvate carboxylase gene family of Sorghum: promoter structu

A; Reference number: JH0381; MUID:91216449; PMID:2022326

A; Accession: JH0381

A; Status: Preliminary

A; Molecule type: DNA

A; Residues: 1-960 <CRE>

A; Cross-references: GB:X55664

C; Superfamily: phosphoenolpyruvate carboxylase

C; Keywords: carbon dioxide fixation; carbon-carbon lyase; phosphoprotein

Query Match Score 42; DB 2; Length 960;

Best Local Similarity 47.6%; Pred. No. 68;

Matches 10; Conservative 1; Mismatches 2; Indels 8; Gaps 1;

Qy 4 LRSNYE-----VKGHRDV 16

Db 850 LRKYETKEKLLQVAGHKDV 870

RESULT 8

S31159

phosphoenolpyruvate carboxylase (EC 4.1.1.31) CP21 - sorghum

C; Species: Sorghum bicolor (sorghum)

C; Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 22-Jun-1999

C; Accession: S31159

C; Genetcs: Lequin, L.; Keryer, E.; Philippe, H.; Gadal, P.; Cretin, C.

Plant Mol. Biol. 21, 487-502, 1993

A; Title: Sorghum phosphoenolpyruvate carboxylase gene family: structure, function and mo

A; Reference number: S31159; MUID:93184205; PMID:8443342

A; Molecule type: DNA

A; Residues: 1-960 <LEP>

A; Cross-references: EMBL:X65137; PID:922614; PMID:CAA46267.1; PID:922615

C; Superfamily: phosphoenolpyruvate carboxylase

C; Keywords: carbon dioxide fixation; carbon-carbon lyase; phosphoprotein

Query Match Score 42; DB 2; Length 960;

Best Local Similarity 47.6%; Pred. No. 68;

Matches 10; Conservative 1; Mismatches 2; Indels 8; Gaps 1;

Qy 4 LRSNYE-----VKGHRDV 16

Db 850 LRKYETKEKLLQVAGHKDV 870

RESULT 9

QYNT

phosphoenolpyruvate carboxylase (EC 4.1.1.31) - common tobacco

C; Species: Nicotiana tabacum (common tobacco)

C; Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 18-Jun-1999

C; Accession: S17440

C; Genetcs: Koizumi, N.; Sato, F.; Terano, Y.; Yamada, Y.

Plant Mol. Biol. 17, 535-539, 1991

A; Title: Sequence analysis of cDNA encoding phosphoenolpyruvate carboxylase from culture

A; Reference number: S17440; MUID:91355949; PMID:1884006

A; Molecule type: mRNA

A; Residues: 1-964 <SYE>

A; Cross-references: EMBL:X59016; PID:g22588; PMID:CAA41758.1; PID:922589

C; Superfamily: phosphoenolpyruvate carboxylase

C; Keywords: carbon dioxide fixation; carbon-carbon lyase; phosphoprotein

Query Match Score 42; DB 2; Length 964;

Best Local Similarity 42.9%; Pred. No. 69;

Matches 9; Conservative 3; Mismatches 1; Indels 8; Gaps 1;

Qy 4 LRSNYE-----VKGHRDV 16

Db 854 LRKYETEETRSLLQIAGHKDL 874

RESULT 10

S25081

phosphoenolpyruvate carboxylase (EC 4.1.1.31) - Flaveria pringlei

S18318

phosphoenolpyruvate carboxylase (EC 4.1.1.31) isoform C4 (clone Ppc1-1) - Flaveria t

C; Species: Flaveria trinervia

C; Date: 22-Nov-1993 #sequence_revision 23-Feb-1996 #text_change 22-Jun-1999

C; Accession: S18318; S60517; S11342

R; Poetsch, W.; Hermans, J.; Westhoff, P.

PEBS Lett. 292, 133-136, 1991

A; Title: Multiple cDNAs of phosphoenolpyruvate carboxylase in the C(4) dicot Flaveri

A; Reference number: S18318; MUID:92070471; PMID:1720398

A; Molecule type: mRNA

A; Residues: 1-966 <POE>

A; Cross-references: EMBL:X61304

A; Accession: S18318

A; Note: only a part of the coding sequence is given in this paper

A; Submitted to the EMBL Data Library, June 1994

A; Reference number: S60517

A; Molecule type: mRNA

A; Residues: 1-256, R / 258-852, 'S', 854-966 <POE>

A; Cross-references: EMBL:X61304; PID:9498698; PMID:CAA3601.1; PID:9498699

A; Note: this is a revision to the sequence from reference S17342

R; Poetsch, W.; Hermans, J.; Westhoff, P.

submitted to the EMBL Data Library,

A; Description: Multiple cDNAs of the phosphoenolpyruvate carboxylase

A; Reference number: S17342

A; Accession: S17342

A; Molecule type: mRNA

A; Residues: 1-180, SMEGFIVWPS', 191, 'MPKTSLLM', 200, 'SRNSMSCKF', 211-852, 'S', 854-966 <E

A; Cross-references: EMBL:X61304

A; Note: this sequence has been revised in reference S60517

C; Superfamily: phosphoenolpyruvate carboxylase

C; Keywords: carbon dioxide fixation; carbon-carbon lyase; phosphoprotein

Query Match Score 42; DB 2; Length 966;

Best Local Similarity 36.4%; Pred. No. 69;

Matches 8; Conservative 5; Mismatches 1; Indels 8; Gaps 1;

Qy 3 FURSNYE-----VKGHRDV 16

Db 853 YLRANYETETKNLLKIAGHKDL 874

RESULT 11

S52853

phosphoenolpyruvate carboxylase (EC 4.1.1.31) - Flaveria pringlei

C; Species: Flaveria pringlei

C; Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 22-Jun-1999

C; Accession: S52853

R; Svensson, P.; Blaesing, O.E.; Westhoff, P.

submitted to the EMBL Data Library, April 1995

A; Description: Characteristics of the orthologous C4 and C3 PPCA phosphoenolpyruvate

A; Reference number: S52853

A; Accession: S52853

A; Status: preliminary

A; Molecule type: mRNA

A; Residues: 1-366 <SYE>

A; Cross-references: EMBL:Z48966; PID:9763096; PMID:CAA88829.1; PID:9763097

C; Superfamily: phosphoenolpyruvate carboxylase

C; Keywords: carbon-carbon lyase; phosphoprotein

Query Match Score 42; DB 2; Length 966;

Best Local Similarity 42.9%; Pred. No. 69;

Matches 9; Conservative 3; Mismatches 1; Indels 8; Gaps 1;

Qy 4 LRSNYE-----VKGHRDV 16

Db 854 LRANYETETKNLLKIAGHKDL 874.

RESULT 12

S25081

phosphoenolpyruvate carboxylase (EC 4.1.1.31) - Flaveria pringlei

C;Species: Flaveria bringlei
 C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 22-Jun-1999
 R;Hermans, J.; Westhoff, P.
 Mol. Gen. Genet. 234: 275-284, 1992
 A;Title: Homologous genes for the C(4) isoform of phosphoenolpyruvate carboxylase in a
 A;Reference number: S25081; PMID:1508152
 A;Status: nucleic acid sequence not shown
 A;Accession: S25081
 A;Residues: 1-967 <HER>
 A;Cross-references: EMBL:X64144; NID:918457; PID:CAA45505.1; PID:918458
 C;Genetics:
 C;Gene: ppA1
 A;Introns: 56/3; 187/2; 215/3; 291/2; 325/3; 355/1; 406/3; 739/3; 868/3
 C;Keywords: phosphoenolpyruvate carboxylase; carbon dioxide fixation; carbon-carbon lyase; phosphoprotein
 C;Keywords: carbon dioxide fixation; carbon-carbon lyase; carboxy-lyase; phosphoprotein

Query Match 48.8%; Score 42; DB 2; Length 967;
 Best Local Similarity 42.9%; Pred. No. 69;
 Matches 9; Conservative 3; Mismatches 1; Indels 8; Gaps 1;
 Qy 4 LRSNYE-----VKGHRDV 16
 Db :|||: |||: 855 LRSNYETKDYLLIKAGHRDL 875

RESULT 13
 B11972 DNA topoisomerase I - Helicobacter pylori (strain J99)
 C;Species: Helicobacter pylori
 A;Variety: strain J99
 C;Accession: B11972
 C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 29-Sep-1999
 R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
 Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
 Nature 397, 176-180, 1999
 A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen
 A;Reference number: A71800; PMID:9923682
 A;Accession: B11972
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-736 <ARN>
 A;Cross references: GB:AE001450; GB:AE001439; NID:94154617; PID:9415462
 A;Experimental source: strain J99
 C;Genetics:
 A;Gene: topA1
 C;Supertfamily: DNA topoisomerase I

Query Match 47.7%; Score 41; DB 2; Length 736;
 Best Local Similarity 55.0%; Pred. No. 76;
 Matches 11; Conservative 2; Mismatches 3; Indels 4; Gaps 2;
 Qy 1 RGFLRSNYEV---KGH-RDV 16
 Db :|||: |||: 17 KNFLDKNTEVASKGHVRD 36

RESULT 14
 D64534 DNA topoisomerase I - Helicobacter pylori (strain 26695)
 C;Species: Helicobacter pylori
 C;Accession: D64534
 C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 29-Sep-1999
 R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenna
 N, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
 Nature 388, 539-541, 1997
 A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpik, P.D.; Smith, H.O.; Fraser, C.
 A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A;Reference number: A64520; PMID:739467
 A;Accession: D64534
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA
 A;Residues: 1-736 <TON>
 A;Cross-references: GB:AE000531; GB:AE000531; NID:92313196; PID:AD07184.1; PID:G231
 C;Superfamily: DNA topoisomerase I
 Query Match 47.7%; Score 41; DB 2; Length 736;
 Best Local Similarity 55.0%; Pred. No. 76;
 Matches 11; Conservative 2; Mismatches 3; Indels 4; Gaps 2;
 Qy 1 RGFLRSNYEV---KGH-RDV 16
 Db :|||: |||: 17 KNFLDKNTEVASKGHVRD 36

RESULT 15
 T01058 hypothetical protein YUPBH12R.42 - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 22-Oct-1999
 C;Accession: T01058
 R:Theologis, A.; Vysotskaya, V.S.; Osborne, B.I.; Federspiel, N.A.; K.
 Oefner, P.; Davis, R.W.
 submitted to the EMBL Data Library, May 1998
 A;Description: Arabidopsis thaliana chromosome 1 YAC YUPBH12R sequence.
 A;Reference number: Z14227
 A;Accession: T01058
 A;Status: translated from GB/EMBL/DDJB
 A;Molecule type: DNA
 A;Cross references: EMBL:AC002986; NID:g2494106; PID:g3152574; PID:9415462
 A;Residues: 1-830 <THE>
 A;Accession: T01058
 A;Molecule type: DNA
 C;Genetics:
 A;Gene: ATSP-YUPBH12R.42
 A;Map Position: 1
 A;Introns: 68/3; 130/3; 160/3; 223/3; 444/3; 482/3; 522/3; 563/3; 596/3; 615/3; 679/2
 C;Superfamily: yeast lanosterol synthase
 Query Match 47.7%; Score 41; DB 2; Length 830;
 Best Local Similarity 50.0%; Pred. No. 86;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 RGFLRSNYEVKGHRDV 16
 Db :|||: |||: 48 RGFLDNFRVKGCSDL 63

Search completed: November 4, 2002, 20:37:55
 Job time : 52 secs

Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	0	Gaps:	0
US-09-462-625-2_COPY_145_160	(1-16) x MMRNATMSI (1-678)		
Qy	1 Arg Gly Phe Leu Arg Ser Asn Tyr Glu Val Lys Gly His Arg Asp Val 16 		
Db	466 CGGGGTTCTTGAGATCACAATGAGTCAGGCAACGGATGTG 513		
RESULT 4			
LOCUS	AF076482 680 bp mRNA linear ROD 15-AUG-1998		
DEFINITION	Mus musculus peptidoglycan recognition protein precursor (Pgrp)		
ORGANISM	Mus musculus		
REFERENCE	AF076482 mRNA, complete cds.		
ACCESSION	AF076482		
VERSION	GI:3342530		
SOURCE			
REMARK			
COMMENT			
AUTHORS	Kang,D., Liu,G., Lundstrom,A., Gelinus,E. and Steiner,H.		
TITLE	A peptidoglycan recognition protein in innate immunity conserved from insects to humans		
JOURNAL	PROC. NATL. ACAD. SCI. U.S.A. 95 (17), 10078-10082 (1998)		
PUBMED	98374308		
PUBLISHER			
ARTICLE	Kang,D., Liu,G., Lundstrom,A., Gelinus,E. and Steiner,H.		
TYPE	Direct Submission		
JOURNAL	Submitted (08-JUL-1998) Microbiology, Stockholm University, Stockholm S-106 91, Sweden		
FEATURES	Location/Qualifiers		
source	1..680		
gene	/organism="Mus musculus" /db_xref="taxon:10090" /tissue_type="spleen"		
CDS	1..680		
product	/product="peptidoglycan recognition protein precursor"		
protein_id	/protein_id="AAC31821."		
db_xref	/db_xref="GI:3342531"		
translation	/translation="MIFACALLALIGLATSCSFIVPRSEWRALPSECSSRLGHPYRYV VISHTAGSFNCNSPDSCSQARNOHYNKNELGWDAYNFJGEDHVEYECRGWNKG DHTGPITWNPMGSGITMGNFMDRVPAKRALRAALNLLECGVSRGLRSNYEVKGHRDV		
sig_peptide	36..584		
mat_peptide	/gene="Pgrp" /gene="Pgrp" /gene="Pgrp" /gene="Pgrp" /gene="Pgrp"		
BASE COUNT	146 a 216 c 174 g 144 t		
ORIGIN			
Alignment Scores:			
Fred. No.:	5.63e-07	Length:	680
Score:	86.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0
US-09-462-625-2_COPY_145_160	(1-16) x AF076482 (1-680)		
Qy	1 Arg Gly Phe Leu Arg Ser Asn Tyr Glu Val Lys Gly His Arg Asp Val 16 		
Db:			
RESULT 5			
LOCUS	BC005582 713 bp mRNA linear ROD 07-AUG-2002		
DEFINITION	Mus musculus, peptidoglycan recognition protein, clone MGC:11430		
ACCESSION	BC005582		
VERSION	BC005582.1		
KEYWORDS	MGC.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	NIH-MGC Project URL: http://mgc.ncbi.nih.gov		
AUTHORS	Strausberg,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-APR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 31A03, Bethesda, MD 20892-2590, USA		
REMARK			
COMMENT			
CONTACT	MGC Help desk		
EMAIL	cgbps>r@mail.nih.gov		
TISSUE PROCUREMENT	Gilbert Smith, Ph.D.		
CDNA LIBRARY PREPARATION	Life Technologies, Inc.		
DNA SEQUENCING CENTER	Baylor College of Medicine Human Genome Sequencing Center		
WEB SITE	http://www.hgsc.bcm.tmc.edu/cdna/		
CONTACT	amg@bcm.tmc.edu		
CONTACT	gunaratne,p,h,garcia,a,m,lulu,x,yale,s,w,hale,s,m,'yoon,v,s,kowis,c,r,richards,s,r,gibbs,r,a.		
CLONE DISTRIBUTION	Clone distribution information can be found through the I.M.A.G.E. Consortium (LNL) at: http://image.llnl.gov		
Series	IRAK Plate: 17 Row: j Column: 22		
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6679292.			
FEATURES			
SOURCE			
Organism	"Mus musculus"		
/db_xref="taxon:10090"			
/map="FB/N"			
/clone="PGC:11430 IMAGE:3969014"			
/tissue_type="Mammary tumor. Metallothionein-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."			
/clone_lib="NCI_CGAP_Mam1"			
/lab_host="DH10B"			
/note="vector: pcMV-SPORT 6"			
47..595			
/codon_start=1			
/product="peptidoglycan recognition protein"			
/protein_id="AAH05821"			
/db_xref="GI:13542756"			
/translaton="MFACALLALIGLATSCSFIVPRSEWRALPSECSSRLGHPYRYV VISHTAGSFNCNSPDSCSQARNOHYNKNELGWDAYNFJGEDHVEYECRGWNKG DHTGPITWNPMGSGITMGNFMDRVPAKRALRAALNLLECGVSRGLRSNYEVKGHRDV			
CDS	1..713		
BASE COUNT	172 a 218 c 178 g 145 t		
ORIGIN			
Alignment Scores:			
Fred. No.:	5.93e-07	Length:	713
Score:	86.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-462-625-2_COPY_145_160 (1-16) × BC005582 (1-713)	QY	1	Arg Gly Phe Leu Arg Ser Asn Tyr Glu Val Lys Gly His Arg Asp Val 16 	RATTUS NORVEGICUS	Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Db	479	CGGGCTCTTGATCCAACTATGAAGTCAAAGGACCCGGATGTG 526	RESULT 6	AF154114 LOCUS AF154114 DEFINITION Rattus norvegicus peptide recognition protein mRNA, complete cds.	REFERENCE 1 (bases 1 to 182897) AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Alsbrooks,S.L., Amaralungue,H.C., Are,J.R., Ayele,M., Banks,T., Barbour,J., Benton,J., Blainage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,K.L., Byrd,N.C., CARRION,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chan,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Barnhart,C., Edgar,D., Edwards,C.C., Elhai,C., Escott,M., Fallis,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabis,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,I.R., Gorrell,J.H., Guevara,W., Gunnarseth,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,M., Hogue,M., Holloway,C., Hollins,B., Homi,I.F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leahy,B., Lewis,J.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Liu,C., Liu,J., Liu,W., Lousegod,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Lung,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,B., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Orlando,R., Pace,A., Payton,B., Perez,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokjan,T., Rolfe,M., Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shoochitari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatsik,A., Tabor,P., Tamersisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,N., Osman,I.K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williamson,G., Williamson,J., Wu,Y., Wu,Y.F., Wu,Y., Wileczky,R., Wooden,S., Worley,K., Weinstock,G., and Gibbs,R.
ACCESSION AF154114.1	VERSION GI:8132325	SOURCE Rattus norvegicus.	FEATURES 1. (bases 1 to 630) ORGANISM Rattus norvegicus. DIRECT SUBMISSION JOURNAL Submitted (25-MAY-1999) Veterinary Comparative Anatomy, Pharmacology, and Physiology, Washington State University, P.O. Box 646520, Pullman, WA 99164, USA FEATURES Location/Qualifiers 1. .630 REFERENCE AUTHORS Lehman,A., Teodecki,E.E. and Krueger,J.M. TITLE /organism="Rattus norvegicus" JOURNAL /strain="Sprague-Dawley" /db-xref="txon:10116" /tissue-type="spineen" gene 1..630 /gene="Pgrp" CDS 20..571 /gene="Pgrp" /note="similar to Homo sapiens and Mus musculus peptide recognition protein; similar to Trichoplusia ni and Mus musculus TAG7" /codon_start=1 /product="peptidoglycan recognition protein PSRP" /protein_id="AAFT3252.1" /db-xref="GI:8132325" /translation="MFIAFWAAPPALIGLADSSCCFVYPRSEENKALPSECISKGUKKKPYRY VVISHTAGSFCSPDSDFOQRNVLQXOMKQLGWCDVAYNLGVGFSEGFRSNSNEVKGRWIK GDHTGPWNPMSTGQLEIYIQWDHYRE" TITLE VOSTLSPGDQLYEIQWDHYRE"	JOURNAL Unpublished REFERENCE 2 (bases 1 to 182897) AUTHORS Worley,K.C.	
BASE COUNT ORIGIN	134 a	193 c 166 g 137 t	COMMENT	TITLE Direct Submission JOURNAL Submitted (16-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA REFERENCE 3 (bases 1 to 182897) AUTHORS Worley,K.C.	
ACCESSION AF110846	VERSION 7	Alignment Scores: Pred. No.: 4.68e-06 Score: 81.00 Percent Similarity: 100.00% Best Local Similarity: 100.00% Query Match: 94.19% DB: 0	TITLE Direct Submission JOURNAL Submitted (17-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA COMMENT On Jul 13, 2002 this sequence version replaced gi:18767308. GENOME CENTER Center: Baylor College of Medicine CENTER CODE: BCM WEB SITE: http://www.hgsc.bcm.tmc.edu/ CONTACT: hgsc-help@bcm.tmc.edu		
DEFINITION Rattus norvegicus clone CH230-105N7, *** SEQUENCING IN PROGRESS	LOCUS AC110846	182897 bp DNA linear HTG 17-JUL-2002 ***, 77 unordered pieces.	PROJECT INFORMATION Center project name: GRPZ Center clone name: CH230-105N7 Sequencing vector: Plasmid; Chemistry: Dye-Terminator Big Dye: 100% of reads Assembly program: Phrap; version 0.990329 Consensus quality: 107348 bases at least Q40 Consensus quality: 113763 bases at least Q30		
KEYWORDS HTGS PHASEI.	SOURCE Norway rat.	ORGANISM Rattus norvegicus			

Consensus quality: 119211 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.edu/docs/Gembank_draft_data.html).
 * NONE: This is a 'working draft' sequence. It currently
 * consists of 77 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

 1 1170: contig of 1170 bp in length
 1 1270: gap of unknown length
 1 2283: contig of 1013 bp in length
 1 2384: 2383: gap of unknown length
 1 3407: contig of 1024 bp in length
 1 3507: 3508: gap of unknown length
 1 4563: 4564: gap of unknown length
 1 4663: 4664: gap of unknown length
 1 5702: 5703: contig of 1039 bp in length
 1 5803: 5802: gap of unknown length
 1 7030: 7029: contig of 1227 bp in length
 1 7129: 7130: gap of unknown length
 1 8304: 8305: contig of 1175 bp in length
 1 8404: 8405: gap of unknown length
 1 9461: 9462: contig of 1057 bp in length
 1 9551: 9552: gap of unknown length
 1 10950: 10951: contig of 1409 bp in length
 1 11071: 11070: gap of unknown length
 1 12728: 12729: contig of 1658 bp in length
 1 12838: 12839: gap of unknown length
 1 14386: 14387: contig of 1558 bp in length
 1 14866: 14867: gap of unknown length
 1 15644: 15645: contig of 1158 bp in length
 1 15744: 15745: gap of unknown length
 1 17208: 17209: contig of 1464 bp in length
 1 17308: 17309: gap of unknown length
 1 18609: 18610: contig of 1301 bp in length
 1 18709: 18710: gap of unknown length
 1 20602: 20603: contig of 1893 bp in length
 1 20702: 20703: gap of unknown length
 1 22104: 22105: contig of 1402 bp in length
 1 22205: 23657: gap of unknown length
 1 23658: 23757: contig of 1453 bp in length
 1 23758: 23759: gap of unknown length
 1 25647: 25648: contig of 1890 bp in length
 1 25748: 25749: gap of unknown length
 1 26812: 26811: contig of 1064 bp in length
 1 26912: 26911: gap of unknown length
 1 28777: 28776: contig of 1866 bp in length
 1 28878: 28877: gap of unknown length
 1 30013: 30012: contig of 1135 bp in length
 1 30113: 30112: gap of unknown length
 1 31475: 31474: contig of 1362 bp in length
 1 33225: 33226: gap of unknown length
 1 33226: 33225: contig of 1651 bp in length
 1 33326: 33325: gap of unknown length
 1 34654: 34653: contig of 1956 bp in length
 1 34754: 34753: gap of unknown length
 1 35799: 35800: contig of 1045 bp in length
 1 35900: 35901: gap of unknown length
 1 37238: 37237: contig of 1339 bp in length
 1 37239: 37238: gap of unknown length
 1 37339: 37338: contig of 1329 bp in length
 1 39394: 39393: gap of unknown length
 1 40730: 40731: contig of 1336 bp in length
 1 40831: 40830: gap of unknown length
 1 43115: 43116: contig of 2285 bp in length
 1 43215: 43216: gap of unknown length
 1 44707: 44708: contig of 1492 bp in length
 1 44807: 44808: gap of unknown length
 1 46334: 46333: contig of 1526 bp in length

46334: gap of unknown length
 46433: gap of unknown length
 48307: 48308: contig of 1874 bp in length
 49700: 49701: contig of 1293 bp in length
 51925: 51926: gap of unknown length
 52025: 52026: contig of 2125 bp in length
 53497: 53026: gap of unknown length
 54498: 54597: contig of 2472 bp in length
 554598: 556275: contig of 1678 bp in length
 56375: 56276: gap of unknown length
 57816: 57817: contig of 1441 bp in length
 57916: 57917: gap of unknown length
 59765: 59766: contig of 1849 bp in length
 59865: 59866: gap of unknown length
 62577: 62578: contig of 2712 bp in length
 62677: 62678: gap of unknown length
 64206: 64207: contig of 1529 bp in length
 64306: 64307: gap of unknown length
 67434: 67435: contig of 3128 bp in length
 67534: 67535: gap of unknown length
 69462: 69463: contig of 1928 bp in length
 69562: 69563: gap of unknown length
 70978: 70979: contig of 1416 bp in length
 71078: 71079: gap of unknown length
 73252: 73253: contig of 2174 bp in length
 73352: 73353: gap of unknown length
 76275: 76276: contig of 2923 bp in length
 76375: 76376: gap of unknown length
 79483: 79484: contig of 3108 bp in length
 79583: 79584: gap of unknown length
 82304: 82305: contig of 2721 bp in length
 82404: 82405: gap of unknown length
 84815: 84816: contig of 2411 bp in length
 84915: 84916: gap of unknown length
 87439: 87440: contig of 2524 bp in length
 87539: 87540: gap of unknown length
 90118: 90119: contig of 2519 bp in length
 90218: 90219: gap of unknown length
 92751: 92752: contig of 2533 bp in length
 92851: 92852: gap of unknown length
 94485: 94486: contig of 1634 bp in length
 94585: 94586: gap of unknown length
 96642: 96643: contig of 2057 bp in length
 96743: 99345: contig of 2603 bp in length

Alignment Scores:
 Pred. No.: 0.00227
 Score: 81.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 94.19%
 DB: 2
 Gaps: 0

US-09-462-625-2_COPY_145_160 (1-16) x AC110846 (1-182897)
 QY 2 GlypheLeuArgSerAsnTyrGluValLysGlyHisArgAspVal 16
 CDR131676 LOCUS Camelus dromedarius mRNA for peptidoglycan recognition protein.
 DEFINITION Camelus dromedarius mRNA for peptidoglycan recognition protein.
 ACCESSION AJ131676
 VERSION AJ131676.1 GR:11990123
 KEYWORDS Peptidoglycan recognition protein.
 SOURCE Camelus dromedarius
 ORGANISM Camelus dromedarius
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.
 REFERENCE 1 (bases 1 to 700)

AUTHORS	Kappeler,S.R., Farah,Z. and Puhan,Z.													
TITLE	Milk as a Source of Camel (Camelus dromedarius) Peptidoglycan													
JOURNAL	Recognition Protein													
REFERENCE	Unpublished													
AUTHORS	Kappeler,S.R.													
TITLE	Direct Submission													
JOURNAL	Submitted (21-DEC-1998) Kappeler S.R., Institute of Food Science, Swiss Federal Institute of Technology, LFO F26, 8092 Zurich, SWITZERLAND													
FEATURES	<table border="1"> <thead> <tr> <th>Location/Qualifiers</th> </tr> </thead> <tbody> <tr> <td>source</td> <td>1..700 /organism="Camelus dromedarius" /sub_species="Somali breed" /db_xref="taxon:9838" /tissue_type="lactating mammary gland" 25..606 /codon_start=1 /product="peptidoglycan recognition protein" /protein_id="CAC19553.1" /db_xref="GI:11990124" /translation="MTRACYLLWALLLISLGAAEDPACGSIVPRREWWRALASEC RERJUTRPVRYVVSHTAGSHCDTASCAQAOANTSYHVRNLGNCDVGNFLIGDGL VYERGQWNKVGAAHGPWTNWLISGMNPKPPRALARQNLLAGCVVALGR SNYEVKGHDRDVQPTLSPGDRLYEIQTWSHYRA", 25..87 /product="peptidoglycan recognition protein" 88..603 /product="peptidoglycan recognition protein" 678..683 /product="peptidoglycan recognition protein"</td> </tr> <tr> <td>sig_peptide</td><td>130 a</td></tr> <tr> <td>mat_peptide</td><td>247 c</td></tr> <tr> <td>polyA_signal</td><td>201 g</td></tr> <tr> <td>BASE COUNT</td><td>122 t</td></tr> <tr> <td>ORIGIN</td><td></td></tr> </tbody> </table>	Location/Qualifiers	source	1..700 /organism="Camelus dromedarius" /sub_species="Somali breed" /db_xref="taxon:9838" /tissue_type="lactating mammary gland" 25..606 /codon_start=1 /product="peptidoglycan recognition protein" /protein_id="CAC19553.1" /db_xref="GI:11990124" /translation="MTRACYLLWALLLISLGAAEDPACGSIVPRREWWRALASEC RERJUTRPVRYVVSHTAGSHCDTASCAQAOANTSYHVRNLGNCDVGNFLIGDGL VYERGQWNKVGAAHGPWTNWLISGMNPKPPRALARQNLLAGCVVALGR SNYEVKGHDRDVQPTLSPGDRLYEIQTWSHYRA", 25..87 /product="peptidoglycan recognition protein" 88..603 /product="peptidoglycan recognition protein" 678..683 /product="peptidoglycan recognition protein"	sig_peptide	130 a	mat_peptide	247 c	polyA_signal	201 g	BASE COUNT	122 t	ORIGIN	
Location/Qualifiers														
source	1..700 /organism="Camelus dromedarius" /sub_species="Somali breed" /db_xref="taxon:9838" /tissue_type="lactating mammary gland" 25..606 /codon_start=1 /product="peptidoglycan recognition protein" /protein_id="CAC19553.1" /db_xref="GI:11990124" /translation="MTRACYLLWALLLISLGAAEDPACGSIVPRREWWRALASEC RERJUTRPVRYVVSHTAGSHCDTASCAQAOANTSYHVRNLGNCDVGNFLIGDGL VYERGQWNKVGAAHGPWTNWLISGMNPKPPRALARQNLLAGCVVALGR SNYEVKGHDRDVQPTLSPGDRLYEIQTWSHYRA", 25..87 /product="peptidoglycan recognition protein" 88..603 /product="peptidoglycan recognition protein" 678..683 /product="peptidoglycan recognition protein"													
sig_peptide	130 a													
mat_peptide	247 c													
polyA_signal	201 g													
BASE COUNT	122 t													
ORIGIN														
Alignment Scores :														
Pred. No.:	0.000178													
Score:	73.00													
Percent Similarity:	93.33%													
Best Local Similarity:	93.33%													
Query Match:	84.88%													
DB:	4													
Gaps:	0													
Oligo:	0													
RESULT	9													
CDR49286	CDB409286													
LOCUS	5358 bp													
DEFINITION	DNA linear MAM 19-DEC-2001													
Camelus dromedarius pgp gene for peptidoglycan recognition protein, flanking region and exons 1-3.														
ACCESION	AJ09286													
VERSION	1													
KEYWORDS	GI:17398982													
SOURCE	peptidoglycan recognition protein: pgp gene.													
ORGANISM	Arabian camel, Camelus dromedarius													
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.														
REFERENCE	1													
AUTHORS	Kappeler,S.R., Farah,Z. and Puhan,Z.													
TITLE	The peptidoglycan recognition protein is expressed in the lactating mammary gland of camels and binds to lactic acid bacteria													
JOURNAL	Unpublished													
REFERENCE	2 (bases 1 to 5358)													
AUTHORS	Kappeler,S.R.													
TITLE	Direct Submission													
JOURNAL	Submitted (19-MAR-2001) Kappeler S.R., Food Science, Swiss Federal Institute of Technology, Zurich, LFO F26, ETHZ, CH-8092 Zurich, SWITZERLAND													
FEATURES	<table border="1"> <thead> <tr> <th>Location/Qualifiers</th> </tr> </thead> <tbody> <tr> <td>source</td> <td>/organism="Camelus dromedarius" /db_xref="taxon:9838"</td> </tr> </tbody> </table>	Location/Qualifiers	source	/organism="Camelus dromedarius" /db_xref="taxon:9838"										
Location/Qualifiers														
source	/organism="Camelus dromedarius" /db_xref="taxon:9838"													
BASE COUNT	1287 a													
ORIGIN	1351 c 1533 g 1187 t													
Alignment Scores :														
Pred. No.:	0.00163													
Score:	73.00													
Length:	5358													
Matches:	14													

A; Reference number: JH0381; MUID:91216449; PMID:2022326

A; Accession: JH0381

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-360 <CRE>

C; Cross-references: GB:X55664

C; Superfamily: Phosphoenolpyruvate carboxylase

C; Keywords: carbon dioxide fixation; carbon-carbon lyase; phosphoprotein

Query Match

48.8%; Score 42; DB 2; Length 960;

Best Local Similarity 47.6%; Pred. No. 68;

Matches 10; Conservative 1; Mismatches 2; Indels 8; Gaps 1;

Qy 4 LRSNYE-----VKGHRDV 16

Db 850 LRKNTETKELLQLQVAGHKDV 870

RESULT 8

S31159

phosphoenolpyruvate carboxylase (EC 4.1.1.31) CP21 - sorghum

C; Species: Sorghum bicolor (sorghum)

C; Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 22-Jun-1999

C; Accession: S31159

R; Lepiniec, L.; Kieber, E.; Philippe, H.; Gadal, P.; Cretin, C.

Plant Mol. Biol. 21, 487-502, 1993

A; Title: Sorghum phosphoenolpyruvate carboxylase gene family: structure, function and mo-

A; Reference number: S31159; MUID:93184205; PMID:8443342

A; Accession: S31159

A; Molecule type: DNA

A; Residues: 1-960 <LEP>

A; Cross-references: EMBL:X65137; NID:922614; PIDN:CAA46267.1; PID:922615

C; Genetics:

A; Introns: 52/3; 183/3; 211/3; 285/3; 320/3; 349/3; 401/3; 734/3; 863/3

C; Superfamily: phosphoenolpyruvate carboxylase; carbon dioxide fixation; carbon-carbon lyase; phosphoprotein

Query Match

48.8%; Score 42; DB 2; Length 960;

Best Local Similarity 47.6%; Pred. No. 68;

Matches 10; Conservative 1; Mismatches 2; Indels 8; Gaps 1;

Qy 4 LRSNYE-----VKGHRDV 16

Db 850 LRKNTETKELLQLQVAGHKDV 870

RESULT 9

QYNT

phosphoenolpyruvate carboxylase (EC 4.1.1.31) - common tobacco

C; Species: Nicotiana tabacum (common tobacco)

C; Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 18-Jun-1999

C; Accession: S17440

R; Koizumi, N.; Sato, F.; Terada, Y.; Yamada, Y.

Plant Mol. Biol. 17, 535-539, 1991

A; Title: Sequence analysis of cDNA encoding phosphoenolpyruvate carboxylase from culture

A; Reference number: S17440; MUID:91355949; PMID:1684006

A; Accession: S17440

A; Molecule type: mRNA

A; Residues: 1-364 <ROI>

A; Cross-references: EMBL:X59016; PIDN:CAA41758.1; PID:g22589

C; Superfamily: phosphoenolpyruvate carboxylase

C; Keywords: carbon dioxide fixation; carbon-carbon lyase; phosphoprotein

Query Match

48.8%; Score 42; DB 1; Length 964;

Best Local Similarity 42.9%; Pred. No. 69;

Matches 9; Conservative 3; Mismatches 1; Indels 8; Gaps 1;

Qy 4 LRSNYE-----VKGHRDV 16

Db 854 LRSNYEETKELLQLQVAGHKDL 874

RESULT 10

RESULT 12

S25081

phosphoenolpyruvate carboxylase (EC 4.1.1.31) - Flaveria pringlei

S18318

phosphoenolpyruvate carboxylase (EC 4.1.1.31) isoform C4 (clone ppc1-1) - Flaveria

C; Species: Flaveria trinervia

C; Date: 22-Nov-1993 #sequence_revision 23-Feb-1996 #text_change 22-Jun-1999

C; Accession: S18318; S60517; S17342

R; Poetsch, W.; Hermans, J.; Westhoff, P.

FEBS Lett. 292, 133-136, 1991

A; Title: Multiple cDNAs of phosphoenolpyruvate carboxylase in the C(4) dicot Flave-

A; Reference number: S18318; MUID:92070471; PMID:1720398

A; Accession: S18318

A; Molecule type: mRNA

A; Residues: 1-256; R'; 258-852; 'S' 854-866 <POE>

A; Cross-references: EMBL:X61304; PIDN:CA43601.1; PID:9498699

A; Note: this is a revision to the sequence from reference S17342

R; Poetsch, W.; Hermans, J.; Westhoff, P.

submitted to the EMBL Data Library, August 1991

A; Description: Multiple cDNAs of the phosphoenolpyruvate carboxylase in the C4 dic-

A; Reference number: S17342

A; Accession: S17342

A; Molecule type: mRNA

A; Residues: 1-180; SMEGFIVWPS' 191; MPKTSILIM' 200; SRNSMSCFK' 211-852; 'S' 854-966

A; Cross-references: EMBL:X61304

A; Note: this sequence has been revised in reference S60517

C; Superfamily: phosphoenolpyruvate carboxylase

C; Keywords: carbon dioxide fixation; carbon-carbon lyase; carboxy-lyase; phosphoprc

Query Match

48.8%; Score 42; DB 2; Length 966;

Best Local Similarity 36.4%; Pred. No. 69;

Matches 8; Conservative 5; Mismatches 1; Indels 8; Gaps 1;

Qy 3 FIRSNYE-----VKGHRDV 16

Db 853 YLRANWEEETKNNLLKAGHKDL 874

RESULT 11

S52853

phosphoenolpyruvate carboxylase (EC 4.1.1.31) - Flaveria pringlei

C; Species: Flaveria pringlei

C; Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 22-Jun-1999

C; Accession: S52853

R; Svensson, P.; Blaesting, O.E.; Westhoff, P.

submitted to the EMBL Data Library, April 1995

A; Description: Characteristics of the orthologous C4 and C3 PPCA phosphoenolpyruvate

A; Reference number: S52853

A; Status: preliminary

A; Molecule type: mRNA

A; Residues: 1-366 <SVE>

A; Cross-references: EMBL:Z48966; NID:9733096; PIDN:CAA8829.1; PID:9763097

C; Superfamily: phosphoenolpyruvate carboxylase

C; Keywords: carbon-carbon lyase

Query Match

48.8%; Score 42; DB 2; Length 966;

Best Local Similarity 42.9%; Pred. No. 69;

Matches 9; Conservative 3; Mismatches 1; Indels 8; Gaps 1;

Qy 4 LRSNYE-----VKGHRDV 16

Db 854 LRANWEEETKNNLLKAGHKDL 874

RESULT 12

S25081

C:Species: Flaveria pringlei
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 22-Jun-1999
 C:Accession: S25081
 R:Hermanns, J.; Westhoff, P.
 Mol. Gen. Genet. 234, 275-284, 1992
 A:Title: Homologous genes for the C(4) isoform of phosphoenolpyruvate carboxylase in a C
 A:Reference number: S25081; MUID:1508152
 A:Accession: S25081
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-967 <HHR>
 A:Cross-references: EMBL:X64144; NID:918457; PIDN:CAA45505.1; PID:918458
 C:Genetics:
 A:Gene: ppca1
 A:Introns: 56/3; 187/2; 215/3; 291/2; 325/3; 355/1; 406/3; 739/3; 868/3
 C:Keywords: Phosphoenolpyruvate carboxylase; carbon dioxide fixation; carbon-carbon lyase; phosphoprotein
 C:Superfamily: Carbon dioxide fixation; carbon-carbon lyase; carboxy-lyase; phosphoprotein

RESULT 15
 T01058 hypothetical protein YUP8H12R.42 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 22-Oct-1999
 C:Accession: T01058
 R:Theologis, A.; Vysotskaya, V.S.; Osborne, B.I.; Federspiel, N.A.; Oehler, P.; Davis, R.W.
 submitted to the EMBL Data Library, May 1998
 A:Description: Arabidopsis thaliana chromosome 1 YAC YUP8H12R sequence.
 A:Reference number: Z14227
 A:Accession: T01058
 A:Status: translated from GB/EMBL/DDJB
 A:Molecule type: DNA
 A:Residues: 1-830 <THE>
 A:Cross references: EMBL:AC002986; NID:92494106; PID:93152574; PIDN: AAC17055.1; GSPD
 C:Genetics:
 A:Map position: 1
 A:Introns: 68/3; 130/3; 160/3; 223/3; 563/3; 596/3; 615/3; 679/
 C:Superfamily: yeast lanosterol synthase
 R:ALM, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Marberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Ives, 397, 176-180, 1999
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen
 A:Reference number: A71800; MUID:99120557; PMID:9923682
 A:Accession: B71972
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-736 <ARN>
 A:Cross-references: GB:AE001450; NID:94154617; PIDN:AAD05695.1; PID:9415462
 A:Experimental source: strain J99
 C:Genetics:
 A:Gene: topA1
 C:Superfamily: DNA topoisomerase I

RESULT 16
 Query Match 48.8%; Score 42; DB 2; Length 957;
 Best Local Similarity 42.9%; Pred. No. 69;
 Matches 9; Conservative 3; Mismatches 1;
 Indels 8; Gaps 1;

Qy 4 LRSNYE-----VKGHRDV 16
 Db 855 LRSNYEETKDKLKIAGHRDL 875

RESULT 17
 B71972 DNA topoisomerase I - Helicobacter pylori (strain J99)
 C:Species: Helicobacter pylori
 A:Variety: strain J99
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 29-Sep-1999
 C:Accession: B71972
 R:ALM, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Marberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Ives, 397, 176-180, 1999
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen
 A:Reference number: A71800; MUID:99120557; PMID:9923682
 A:Accession: B71972
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-736 <ARN>
 A:Cross-references: GB:AE001439; NID:94154617; PIDN:AAD05695.1; PID:9415462
 A:Experimental source: strain J99
 C:Genetics:
 A:Gene: topA1
 C:Superfamily: DNA topoisomerase I

RESULT 18
 Query Match 47.7%; Score 41; DB 2; Length 736;
 Best Local Similarity 55.0%; Pred. No. 76;
 Matches 11; Conservative 2; Mismatches 2;
 Indels 4; Gaps 2;

Qy 1 RGFLRSNYEV---KGHRDV 16
 Db 17 KNFLDKNVEVASKGHYRDL 36

Search completed: November 4, 2002, 20:37:55
 Job time : 52 secs

D64534 DNA topoisomerase I - Helicobacter pylori (strain 266695)
 C:Species: Helicobacter pylori
 C:Accession: D64534
 C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 29-Sep-1999
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, K.; Khalak, H.G.; Glodek, A.; McKenna, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpov, P.D.; Smith, H.O.; Fraser, C.
 A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A:Reference number: A64520; MUID:97394467; PMID:9252185
 A:Accession: D64534
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown

Result No.	Score	Query	Match	Length	DB ID	Description
1	86	100.0	182	1	PGRP_MOUSE	088593 mus musculus
2	64	74.4	196	1	PGRP_HUMAN	075594 homo sapiens
3	48	55.8	182	1	PGRP_TRINI	076537 trichoplusia
4	41	50.0	960	1	CAP2_SORBI	P29194 sorghum bic
5	43	50.0	967	1	CAP2_MATZE	P51059 zea mays (m)
6	42	48.8	960	1	CAP1_SORBI	P29195 sorghum bic
7	42	48.8	964	1	CAPP_TOBAC	P227154 nicotiana t
8	42	48.8	967	1	CAPP_TOBAC	P01647 flaveria pr
9	41	47.7	123	1	ANG2_BOVIN	P80329 bos taurus
10	41	47.7	447	1	VIPR_CARAU	Q90308 carassius a
11	41	47.7	736	1	TOP1_HELPY	P55391 helicobacter
12	41	47.7	1065	1	KDG1_HUMAN	Q75312 homo sapiens
13	41	46.5	806	1	LON_BORBU	Q593185 borrelia bu
14	40	46.5	809	1	PUPB_PSEPU	P38047 pseudomonas
15	40	46.5	819	1	LON_CHLMP	Q99K50 chlamydia m
16	40	46.5	819	1	LON_CHLMP	Q8448 chlamydia p
17	40	46.5	819	1	LON_CHLTR	P04090 mesembryant
18	40	46.5	966	1	CAP1_MESCR	P29193 schizosacch
19	40	46.5	966	1	CAP1_SACHY	Q04836 rattus norvegicus
20	40	46.5	1025	1	THDE_RAT	Q53352 methanococcus
21	40	46.5	1122	1	Y942_MEJU	Q51558 borrelia bu
22	39	45.3	651	1	Y942_MEJU	P30694 flaveria tr
23	39	45.3	813	1	LON2_BORBU	Q02730 flaveria au
24	39	45.3	966	1	CAP2_FLCTR	Q01648 flaveria tr
25	39	45.3	966	1	CAP2_FLAU	P49896 rana catesbeiana
26	39	45.3	967	1	CAP1_FLCTR	P17013 halobacteri
27	38	44.2	264	1	IOD2_RANCA	Q9ukus homo sapiens
28	38	44.2	370	1	YB11_HUMAN	Q8446 salmonella
29	38	44.2	1024	1	THDE_HUMAN	Q181_MEJU
30	37.5	43.6	340	1	PURR_SALTY	Q57640 methanococcus
31	37	43.0	246	1	Y181_MEJU	Q34827 bacillus su
32	37	43.0	293	1	YKDM_BACUS	P42504 rhodobacter
33	37	43.0	332	1	HEM2_RHOC	

Scoring table:	BLOSUM62
GappP	10.0 , gapext 0.5
Searched:	112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters:	112892
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing:	Minimum Match 0*
	Maximum Match 100*
	Listing first 45 summaries
Database :	SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	86	100.0	182	1	PGRP_MOUSE	088593 mus musculus
2	64	74.4	196	1	PGRP_HUMAN	075594 homo sapiens
3	48	55.8	182	1	PGRP_TRINI	076537 trichoplusia
4	41	50.0	960	1	CAP2_SORBI	P29194 sorghum bic
5	43	50.0	967	1	CAP2_MATZE	P51059 zea mays (m)
6	42	48.8	960	1	CAP1_SORBI	P29195 sorghum bic
7	42	48.8	964	1	CAPP_TOBAC	P227154 nicotiana t
8	42	48.8	967	1	CAPP_TOBAC	P01647 flaveria pr
9	41	47.7	123	1	ANG2_BOVIN	P80329 bos taurus
10	41	47.7	447	1	VIPR_CARAU	Q90308 carassius a
11	41	47.7	736	1	TOP1_HELPY	P55391 helicobacter
12	41	47.7	1065	1	KDG1_HUMAN	Q75312 homo sapiens
13	41	46.5	806	1	LON_BORBU	Q593185 borrelia bu
14	40	46.5	809	1	PUPB_PSEPU	P38047 pseudomonas
15	40	46.5	819	1	LON_CHLMP	Q99K50 chlamydia m
16	40	46.5	819	1	LON_CHLTR	Q8448 chlamydia p
17	40	46.5	819	1	CAP1_MESCR	P04090 mesembryant
18	40	46.5	966	1	CAP1_SACHY	P29193 schizosacch
19	40	46.5	966	1	THDE_RAT	Q04836 rattus norvegicus
20	40	46.5	1025	1	IOD2_RANCA	Q53352 methanococcus
21	40	46.5	1122	1	Y942_MEJU	Q51558 borrelia bu
22	39	45.3	651	1	Y942_MEJU	P30694 flaveria tr
23	39	45.3	813	1	LON2_BORBU	Q02730 flaveria au
24	39	45.3	966	1	CAP2_FLCTR	Q01648 flaveria tr
25	39	45.3	966	1	CAP2_FLAU	P49896 rana catesbeiana
26	39	45.3	967	1	CAP1_FLCTR	P17013 halobacteri
27	38	44.2	264	1	IOD2_RANCA	Q9ukus homo sapiens
28	38	44.2	370	1	YB11_HUMAN	Q8446 salmonella
29	38	44.2	1024	1	THDE_HUMAN	Q181_MEJU
30	37.5	43.6	340	1	PURR_SALTY	Q57640 methanococcus
31	37	43.0	246	1	Y181_MEJU	Q34827 bacillus su
32	37	43.0	293	1	YKDM_BACUS	P42504 rhodobacter
33	37	43.0	332	1	HEM2_RHOC	

RESULT 1						
ID	PGRP_MOUSE	STANDARD	PRT;	182 AA.		
AC	088593; 062185;					
DT	16-OCT-2001 (Rel. 40, Created)					
DT	16-OCT-2001 (Rel. 40, Last sequence update)					
DE	Peptidoglycan recognition protein precursor (Cytokine tag7).					
GN	PGLYRP OR PGRP OR TAG7.					
OS	Mus musculus (Mouse)					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.					
OX	NCBI_TaxID=10090;					
RN	[1]					
RP	SEQUENCE FROM N.A., FUNCTION AND TISSUE SPECIFICITY.					
RC	TISSUE=Spleen;					
RX	Medline-#98374308; PubMed-9707603;					
RA	Kang D., Liu G., Lundstrom A., Gelinus E., Steiner H.;					
RA	"A peptidoglycan recognition protein in innate immunity conserved from insects to humans."					
RL	Proc. Natl. Acad. Sci. U.S.A. 95:10078-10082(1998).					
RN	[2]					
RP	SEQUENCE FROM N.A., FUNCTION AND SUBCELLULAR LOCATION.					
RX	Medline-#98325081; PubMed-9660837;					
RA	Kiselev S.L., Rustikova O.S., Korobko E.V., Prokhorchouk E.B., Kabishhev A.A., Lukandin E.M., Georgiev G.P.;					
RA	"Molecular cloning and characterization of the mouse tag7 gene encoding a novel cytokine."					
RL	J. Biol. Chem. 273:18633-18639(1998).					
RN	[3]					
RP	SEQUENCE FROM N.A.					
RA	Slayton W.B., Riggs A., Hancock J.D., Zaugg J.K., Le T.V., Trautman M.S., Spangler G.J., Carroll W.L., Schibler K.R.;					
RA	"Granulocyte-colony stimulating factor up-regulates expression of murine tag7 during myeloid differentiation."					
RL	Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.					
RN	[4]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=C57BL/6J; TISSUE=Small intestine;					
RX	Medline-#21085660; PubMed-11217851;					
RA	Kawai J., Shinagawa A., Shibusawa K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Miyosawa H., Kondo S., Yamamoto I., Saito T., Oizaki Y., Gojobori T., Tomita T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batzloff S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kiehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Wagner L., Washio T., Schriml L.M., Stabili F., Suzuki R., Tomita M., Wagner L., Quackenbush J., Blake J., Borrelli D., Fletcher C., Fujita M., Gariboldi M., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sesaki H., Sato K., Schoenbach C., Soya T., Shibusawa Y., Storch K.-F., Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whiting L.,					

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohtsuki S.,
 RA Hayashizaki Y.; "Functional annotation of a full-length mouse cDNA collection.",
 RL Nature 409:685-690(2001).
 RN [5]
 SEQUENCE FROM N.A.
 RP Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: BINDS SPECIFICALLY TO PEPTIDOGLYCAN AND IS INVOLVED IN
 CC INNATE IMMUNITY. THE SOLUBLE FORM TRIGGERS APOPTOSIS IN VITRO.
 CC -!- ASSOCIATED LOCATION: EXISTS IN BOTH SOLUBLE AND MEMBRANE-
 CC -!- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN SPLEEN AND LUNG. ALSO
 CC DETECTED IN BRAIN AND THYMUS. IN THE LUNG, EXPRESSED IN THE
 CC INTERALVEOLAR SPACE, IN THE BRAIN, EXPRESSED IN THE PURKINJE CELLS
 CC OF THE CEREBELLUM AND IN CERTAIN LAYERS OF NEURONS IN THE
 CC HIPPOCAMPUS. ALSO DETECTED IN CELLS FILLING THE SPACE WITHIN THE
 CC INTESTINAL VILLUS.
 CC -!- SIMILARITY: BELONGS TO THE PEPTIDOGLYCAN RECOGNITION PROTEIN
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR AF076482; AAC31821; 1;
 DR AF19843; AAF06335; 1;
 DR AK008335; BAB25611; 1;
 DR BC005582; AAH05582; 1;
 DR EMBL; X83374; CAA6133.; ALT_SEQ.
 DR MGD; MGI:1345092; Pglyrp.
 KW Immune response; Cytokine; Apoptosis; Signal.
 FT SIGNAL 1 18 PEPTIDOGLYCAN RECOGNITION PROTEIN.
 FT CHAIN 19 182 PEPTIDOGLYCAN RECOGNITION PROTEIN.
 FT DISULFID 54 60 POTENTIAL.
 SQ SEQUENCE 182 AA; 20489 MW; 9844E2137F04 CRC64;
 Query Match 100.0%; Score 86; DB 1; Length 182;
 Best Local Similarity 100.0%; Pred. No. 4.e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 RGFLRSNYEVKGHRDV 16
 Db 145 RGFLRSNYEVKGHRDV 160

RESULT 2

ID	PGRP	HUMAN	STANDARD	PRT	196 AA.
AC	075594;				
DT	16-OCT-2001	(Rel. 40, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DE		Peptidoglycan recognition protein precursor (SBB168).			
GN		PGLYRP OR PGRP.			
OS		Homo sapiens (Human).			
OC		Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OX		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
NCBI_TAXID		9606;			
OX					

RN [1] SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RP TISSUE=Bone marrow;
 RC MEDLINE=3874308; PubMed=9707603;
 RA Kang D.; Liu G.; Lundstrom A.; Gelius E.; Steiner H.;
 RT "A peptidoglycan recognition protein in innate immunity conserved from
 RL insects to humans.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:10078-10082(1998).

RX MEDLINE=98374308; PubMed=9707603;
 RA Kang D., Liu G., Lundstroem A., Geilius E., Steiner H.;
 RT "A peptidoglycan recognition protein in innate immunity conserved from
 insects to humans.";
 RT Proc. Natl. Acad. Sci. U.S.A. 95:10078-10082(1998).
 RL -!
 CC FUNCTION: BINDS SPECIFICALLY TO PEPTIDOGLYCAN AND TRIGGERS THE
 PROTEINOXYDASE CASCADE WHICH IS AN IMPORTANT INNATE IMMUNE
 DEFENSE MECHANISM.
 CC -!- SUBUNIT: MONOMER (PROBABLE).
 CC -!- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN FAT BODY WITH WEAK
 EXPRESSION OBSERVED IN HEMOCYTE. NO EXPRESSION DETECTED IN
 GUT.
 CC -!- INDUCTION: BY BACTERIAL CHALLENGE.
 CC -!- SIMILARITY: BELONGS TO THE PEPTIDOGLYCAN RECOGNITION PROTEIN
 FAMILY.
 CC -!
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -!
 DR EMBL; AF076481; AAC31820.1; -.
 KW Immune response; Signal.
 FT SIGNAL 1 16 PEPTIDOGLYCAN RECOGNITION PROTEIN.
 FT CHAIN 1 7 182 PEPTIDOGLYCAN RECOGNITION PROTEIN.
 FT DISUFTID 18 140 POTENTIAL.
 FT DISUFTID 54 60 POTENTIAL.
 SQ SEQUENCE 182 AA; 20572 MW; 5663.E76.2AE34794 CRC64;
 Query Match 55 8%; Score 48; DB 1; Length 182;
 Best Local Similarity 50.0%; Pred. No. 0.23; Mismatches 3; Indels 0; Gaps 0;
 Matches 8; Conservative 3; Standard PRT; 960 AA.
 QY 1 RGFLRSNYEKVKGHRDV 16
 Db 144 RGHTANVHVGHRDL 159
 RN 1
 RESULT 4 CAP2_SORBI
 ID CAP2_SORBI STANDARD; PRT; 960 AA.
 AC P29134;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DS Phosphoenolpyruvate carboxylase 2 (EC 4.1.1.31) (PEPCASE) (CP28).
 OS Sorghum bicolor (Sorghum vulgare).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Sorghum.
 NCBI_TaxID=4558;
 RN 1
 RP SEQUENCE FROM N.A.
 RX MEDLINE=12032766; PubMed=1840686;
 RA Lepiniec L., Santi S., Kermer E., Amiet V., Vidal J., Gadal P.,
 RA Cretin C.;
 RT "Complete nucleotide sequence of one member of the Sorghum
 phosphoenolpyruvate carboxylase gene family.";
 RL Plant Mol. Biol. 17:1077-1079(1991).
 -!- FUNCTION: TO FORM OXALOACETATE, A FOUR-CARBON DICARBOXYLIC ACID
 SOURCE FOR THE TRICARBOXYLIC ACID CYCLE.
 CC -!- CATALYTIC ACTIVITY: Phosphate + oxaloacetate = H(2)O +
 CC Phosphoenolpyruvate + CO(2).
 CC -!- PATHWAY: TRICARBOXYLIC ACID CYCLE. THIS ISOZYME IS INVOLVED IN
 CC C3 PHOTOSYNTHESIS.
 CC -!- SUBUNIT: HOMOTRIMER.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE PEPCASE FAMILY.
 CC -!
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

DR	HSSP; P00864; 1FIY.	DR	Pfam; PF00311; PEPCase; 1.
DR	Pfam; IPR01449; PEPCase.	DR	PRINTS; PR00311; PEPCASE.
DR	PRINTS; PR00310; PEPCARBOXLASE.	DR	PROSITE; PS00293; PEPCASE_2; 1.
DR	PROSITE; PS00393; PEPCASE_2; 1.	DR	PROSITE; PS00781; PEPCASE_1; 1.
DR	PROSITE; PS00781; PEPCASE_1; 1.	KW	Lyase; Carbon dioxide fixation; Allosteric enzyme; Multigene family;
KW	Lyse; Carbon dioxide fixation; Allosteric enzyme; Multigene family;	KW	Tricarboxylic acid cycle; Phosphorylation; Photosynthesis.
KW	Tricarboxylic acid cycle; Photosynthesis; Phosphorylation.	FT	MOL_RES 11 BY SIMILARITY.
FT	MOD_RES 11 BY SIMILARITY.	FT	ACT_SITE 1172 BY SIMILARITY.
FT	ACT_SITE 1172 BY SIMILARITY.	FT	ACT_SITE 172 BY SIMILARITY.
FT	ACT_SITE 172 BY SIMILARITY.	FT	CONFFLICT 291 292 KH -> N (IN REF. 2).
SQ	SEQUENCE 964 AA; 110163 MW; 5590F55694B22DE4 CRC64;	SQ	SEQUENCE 967 AA; 110628 MW; 8317EA8078C038C0 CRC64;
Query Match	48.8%; Score 42; DB 1; Length 967;	Query Match	48.8%; Score 42; DB 1; Length 967;
Best Local Similarity	42.9%; Pred. No. 18;	Best Local Similarity	42.9%; Pred. No. 18;
Matches 9; Conservative 3; Mismatches 1; Indels 8; Gaps 1;	Matches 9; Conservative 3; Mismatches 1; Indels 8; Gaps 1;		
Qy	4 LRSNYE-----VKGHRDV 16	Qy	4 LRSNYE-----VKGHRDV 16
; ;	; ;	Db	855 LRANYEETKDYLLKIAGRDL 875
RESULT 8		RESULT 9	
CAPI_FLAAPR	ID CAPI_FLAAPR STANDARD; PRT; 967 AA.	ANG2_BOVIN	STANDARD; PRT; 123 AA.
ID	Q01647;	ID	ANG2_BOVIN
AC	01-APR-1993 (Rel. 25, Created)	AC	P80929;
DT	01-APR-1993 (Rel. 25, Last sequence update)	DT	01-NOV-1997 (Rel. 35, Created)
DT	15-JUL-1999 (Rel. 38, Last annotation update)	DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	PhosphoenolPyruvate carboxylase (EC 4.1.1.31) (PEPCASE).	DE	Angiopeptin-2 (EC 3.1.27.-).
GN	PPCA.	GN	ANG2.
OS	Bos taurus (Bovine).	OS	Bos taurus (Bovine).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC	Eukaryota; Metazoa; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bos.	OC	Bovidae; Bovinae; Bos.
NCBI_TAXID=9913;		NCBI_TAXID=9913;	
RN	[1]	RN	[1]
RP	SEQUENCE FROM N.A.	RP	SEQUENCE-Serum, and Milk,
RX	MEDLINE=92374996; PubMed=1508152;	RX	MEDLINE=97410980; PubMed=9266695;
RA	Strydom D.J., Bond M.D., Vallee B.L.;	RA	Strydom D.J., Bond M.D., Vallee B.L.;
RA	"Homologous genes for the C4 isoform of phosphoenolpyruvate	RT	"An angiogenic protein from bovine serum and milk -- purification and
RT	carboxylase in a C3 and a C4 Flavera species.";	RT	primary structure of angiogenin-2."
CC	Mol. Genet. 234:275-284(1992).	RL	Eur. J. Biochem. 247:535-544(1997).
CC	[2]	CC	-1- FUNCTION: BINDS TIGHTLY TO PLACENTAL RIBONUCLEASE INHIBITOR AND
CC	Flavera pringlei.	CC	HAS VERY LOW RIBONUCLEASE ACTIVITY. HAS POTENT ANGIOGENIC
CC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	CC	ACTIVITY. ANGIOGENIN INDUCES VASCULARIZATION OF NORMAL AND
CC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	CC	MALIGNANT TISSUES. ABOLISHES PROTEIN SYNTHESIS BY SPECIFICALLY
CC	Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;	CC	-1- HYDROLYZING CELLULAR TRNS.
CC	Heliantheae; Flaveriae.	CC	-1- TISSUE SPECIFICITY: SERUM, AND MILK.
CC	NCBI_TAXID=4226;	CC	-1- SIMILARITY: BELONGS TO THE PANCREATIC RIBONUCLEASE FAMILY.
RN	[1]	CC	DR P1012; IAGI.
RP	SEQUENCE FROM N.A.	CC	DR Interpro; IPR01427; RNaseA.
RC	TISSUE=Leaf;	CC	DR Pfam; PF00074; rnase; 1.
RA	Svensson P., Blaesing O.E., Westhoff P.;	CC	DR Prints; PD00053; rnase; 1.
RA	"Homologous genes for the C4 isoform of phosphoenolpyruvate	CC	DR ProDom; SM00092; rnase; 1.
RL	SOURCE FOR THE TRICARBOXYLIC ACID CYCLE.	CC	DR SMART; SM00092; rnase_pc; 1.
CC	-1- FUNCTION: TO FORM OXALOACETATE, A FOUR-CARBON DICARBOXYLIC ACID	DR	DR P0012; RNase; Endonuclease; Angiogenesis; 1.
CC	-1- CATALYTIC ACTIVITY: Phosphate + oxaloacetate = H(2)O +	KW	KW Hydrolease; Nuclease; Endonuclease; Glycoprotein.
CC	CC Phosphoenolpyruvate + CO(2).	FT	FT PROTEIN SYNTHESIS INHIBITOR; GLYCOPROTEIN.
CC	-1- ENZYME REGULATION: BY LIGHT-REVERSIBLE PHOSPHORYLATION.	FT	FT PYRROLIDONE CARBOXYLIC ACID.
CC	CC PATHWAY: TRICARBOXYLIC ACID CYCLE. THIS ISOZYME IS INVOLVED IN	FT	FT ACT_SITE 12 BY SIMILARITY.
CC	C3 PHOTOSYNTHESIS	FT	FT ACT_SITE 39 BY SIMILARITY.
CC	-1- SUBUNIT: HOMOTETRAMER.	FT	FT DISULFID 113 BY SIMILARITY.
CC	-1- SIMILARITY: BELONGS TO THE PEPCASE FAMILY.	FT	FT DISULFID 25 80 BY SIMILARITY.
CC	CC This SWISS-PROT entry is copyright. It is produced through a collaboration	FT	FT DISULFID 38 91 BY SIMILARITY.
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	FT	FT DISULFID 56 106 BY SIMILARITY.
CC	the European Bioinformatics Institute. There are no restrictions on its	FT	FT CARBOHYD 33 33 N-LINKED (GLCNAC).
CC	use by non-profit institutions as long as its content is in no way	SQ	SQ SEQUENCE 123 AA; 14522 MW; B703B9839919FD2F CRC64;
CC	modified and this statement is not removed. Usage by and for commercial	DR	DR Z48966; CAA8829.1; Score 41; DB 1; Length 123;
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/)	DR	DR S25081; S25081; Best Local Similarity 52.9%; Pred. No. 2.6; Gaps 1;
CC	or send an email to license@isb-sib.ch).	DR	DR HSSP; P00864; 1FIY;保守性 2; Mismatches 4; Indels 2; Gaps 1;
DR	InterPro; IPR01449; PEPCase.		

FT ZN_FING 663 689 C4-TYPE 3.
 FT ZN_FING 702 725 C4-TYPE 4.
 FT ACT_SITE 297 297 DNA CLEAVAGE (BY SIMILARITY).
 .SQ SEQUENCE 736 AA; 83262 MW; 779EF655F3C14CF CRC64;

Query Match Score: 41; DB 1; Length 736;
 Best Local Similarity 55.0%; Pred. No. 19;
 Matches 11; Conservative 2; Mismatches 3; Indels 4; Gaps 2;

RESULT 12 ID TOP1_HELPY PRT; 736 AA.
 ID TOP1_HELPY STANDARD; PRT; 736 AA.
 AC P55991;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA topoisomerase I (EC 5.99.1.2) (Omega-protein) (Relaxing enzyme)
 DE (Untwisting enzyme) (Swivelase).
 GN TPA OR HP0116.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OX NCBI_TAXID=210;
 RN [1];
 RP SEQUENCE FROM N.A.
 RX STRAIN=973945 / ATCC 700392;
 RX MEDLINE=973945/; PubMed=9252185;
 RA Tomb J.-F., White A.R., Kerlavage A.R., Clayton R.A., Sutton G.G., Gill S., Dougherty B.A., Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A., Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S., Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A., McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K., Berg D.E., Gooley J.D., Utterback T.R., Petersen J.D., Kelley J.M., Cotton M.D., Weldman J.M., Fujii C., Bowman C., Wathey L., Wallin E., Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M., Venter J.C.;
 RT "The complete genome sequence of the gastric pathogen Helicobacter pylori."
 RL Nature 388:539-547(1997);
 CC -I- FUNCTION: THE REACTION CATALYZED BY TOPOISOMERASES LEADS TO THE CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER.
 CC -I- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded DNA, followed by passage and rejoicing.
 CC -I- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -I- MISCELLANEOUS: WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA BACKBONE BOND, IT SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN WHICH A TYROSYL OXYGEN IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS AT ONE END OF THE ENZYME-SEVERED DNA STRAND.
 CC -I- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@lsb-sib.ch).

CC EMBL; AE000531; AAD07184.1; -
 DR HSSP; P06612; 1ECL.
 DR TIGR; HP0116; -
 DR InterPro; IPR02936; DNAPrim_toprinn.
 DR InterPro; IPR03601; DNATop1_ATP_bind.
 DR InterPro; IPR03602; DNATop1_DNA_bind.
 DR InterPro; IPR00380; Prok_topisomerase.
 DR Pfam; PF01131; Topoisom_bac; 1.
 DR Pfam; PF01396; zf-C4_TopoIsom; 3.

DR Pfam; PF01751; Toprinn; 1.
 DR PRINTS; PR00417; PRTPISOMASE1.
 DR SMART; SMO0437; TOP1AC; 1.
 DR SMART; SMO0446; TOP1BC; 1.
 DR TIGRFAMS; TIGR01051; topa_bact; 1.
 DR PROSITE; PS00396; TOPOTISOMERASE_I_PROK; 1.
 KW Isomerase; Topoisomerase; DNA-binding; Zinc-finger; Metal-binding; Repeat; Complete proteome.
 KW Repeat; Complete proteome.
 FT ZN_FING 572 598 C4-TYPE 1.
 FT ZN_FING 616 642 C4-TYPE 2.
 FT ZN_FING 663 689 C4-TYPE 3.
 FT ZN_FING 702 725 C4-TYPE 4.
 FT ACT_SITE 297 297 DNA CLEAVAGE (BY SIMILARITY).
 SQ SEQUENCE 736 AA; 83195 MW; 357CE39735F48475 CRC64;
 Query Match Score: 47.7%; DB 1; Length 736;
 Best Local Similarity 55.0%; Pred. No. 19;
 Matches 11; Conservative 2; Mismatches 3; Indels 4; Gaps 2;

Qy 1 RGFLRSNEYEV---KGH-RDV 16
 Db 17 KNFLDKNEYEVASKGHYRDL 36

RESULT 13 ID RDGL_HUMAN PRT; 1065 AA.
 ID RDGL_HUMAN STANDARD; PRT; 1065 AA.
 AC 075912; Q9NZ49;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Diacylglycerol kinase, iota (EC 2.7.1.107) (Diacylglyceride kinase) (DGK-iota) (DAG kinase iota).
 DE DGKI.
 GN Human sapiens (Human).
 OS Bokaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OC NCBI_TAXID=9606;
 RN [1];
 RP SEQUENCE FROM N.A.
 RX TISSUE=Retina;
 RX MEDLINE=99047655; PubMed=9830018;
 RA Ding L., Traer E., McIntyre T.M., Zimmerman G.A., Prescott S.M.; RT The cloning and characterization of a novel human diacylglycerol kinase, DGK-iota.; RL J. Biol. Chem. 273:32746-32752(1998).
 RN RP SEQUENCE OF 135-1065 FROM N.A. AND VARIANT PHE-153.
 RX MEDLINE=20173854; PubMed=10705894;
 RA Bowe S.J., Sullivan L.S., Ding L., Traer E., Prescott S.M., Birch D.G., Kennan A., Humphries P., Daiger S.P., RT Evaluation of human diacylglycerol kinase iota, DGKI, a homolog of Drosophila rdgA, in inherited retinopathy mapping to 7q.; RL Mol. Vision 6:6-9(2000).
 CC -I- CATALYTIC ACTIVITY: ATP + 1,2-diacylglycerol = ADP + 1,2-diacylglycerol-3-phosphate.
 CC -I- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
 CC -I- SIMILARITY: BELONGS TO THE EUKARYOTIC DIACYLGLYCEROL KINASE FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG BINDING DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 RANK REPEATS.
 CC -I- SIMILARITY: CONTAINS 1 MARKS HOMOLOGY REGION.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@lsb-sib.ch).

Sequence Comparison Report									
Query Sequence		Target Sequence							
Best Local Similarity					Score 41; DB 1; Length 1065;				
Pred. No.					29;				
Smart ID					SMM0382; AAA; 1.				
DR					SMART; SMM0464; LON; 1.				
DR					DR				
DR					PRINMS; PRO00830; ENDOLAPTAZE.				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				
DR					DR				

DR TIGRFAMS; TIGR00763; lon; 1;
 DR PROSITE; PS01046; LON_SER; 1.
 KW Hydrolase; Serine protease; ATP-binding: Complete proteome.
 FT NP BIND 389 396 ATP (POTENTIAL).
 FT ACT_SITE 714 714 SIMILARITY
 SQ SEQUENCE 806 AA: 90709 MW: 71F488D719B1424E CRC64;

Query Match 46.5%; Score 40; DB 1; Length 806;
 Best Local Similarity 53.8%; Pred. No. 32;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 Qy 2 GFLRSNYEVKGR 14
 Db 416 GGMRDESEIKGHR 428

RESULT 15
 PUPB_PSEPU PUPB_PSEPU STANDARD; PRT; 809 AA.
 ID P38017; AC P38017; DT 01-OCT-1994 (Rel. 30, Created)
 RX MEDLINE=93316856; PubMed=8392140;
 RC STRAIN=WCS338;
 RX RA Koster M., van de Vossenberg J., Leong J., Weisbeek P.J.;
 RT "Identification and characterization of the pupB gene encoding an
 inducible ferric-pseudobactin receptor of *Pseudomonas putida*
 WCS338."
 RT Mol. Microbiol. 8:591-601 (1993).
 RL -I- FUNCTION: SPECIFIC RECEPTOR FOR THE SIDEROPHORES FERRIC
 CC PSEUDOBACTIN BN8 AND FERRIC PSEUDOBACTIN BN7. IRON CHELATING
 CC MOLECULES THAT ALLOW THE ORGANISM TO EXTRACT IRON FROM THE
 CC ENVIRONMENT, SPECIALLY UNDER IRON-RESTRICTIVE CONDITIONS.
 CC -I- SUBCELLULAR LOCATION: Outer membrane.
 CC -I- INDUCTION: BY IRON LIMITATION IN ADDITION TO THE PRESENCE OF ONE
 CC OF THE TWO COGNATE PSEUDOBACTINS BN8 OR BN7.
 CC -I- SIMILARITY: LOCAL TO OTHER TONE-DEPENDENT RECEPTOR PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).

CC DR EMBL; X73598; CAA51995.1;
 DR PIR; S32899; S32899.
 DR InterPro; IPR00531; TonB_boxC.
 DR Pfam; PF00593; TonB_boxC_1.
 DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; FALSE_NEG.
 KW Outer membrane; Iron transport; Transport; Signal; Receptor.
 FT SIGNAL 1 45 POTENTIAL.
 FT CHAIN 46 809 FERRIC-PSEUDOBACTIN BN7/BN8 RECEPTOR.
 FT DOMAIN 156 160 POLY SER.
 FT SITE 792 809 TONB C-TERMINAL BOX.
 SQ SEQUENCE 809 AA: 88389 MW: 03339F6E788A8C0D CRC64;

Query Match 46.5%; Score 40; DB 1; Length 809;
 Best Local Similarity 66.7%; Pred. No. 33;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 RGFLRSNYEVKGR

DR ||||| ||||| |||||
 DR 217 RGFAIQNYEVDG 228

Search completed: November 4, 2002, 20:35:04
 Job time : 38 secs

Copyright (c) 1993 - 2002 Compugen Ltd.

OM_protein - protein search, using sw model

Run on : November 4, 2002, 19:54:31 ; Search time 94 Seconds
 Perfect score: 86

Sequence: 1 RGFLRSNYEVKGHRDV 16

Scoring table: BLOSUM62

Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTRMBL_21:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_nthic:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rabbit:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriaph:*

17: sp_archeap:*

17	44	51.2	1764	2	Q93t34 haemophilus
	18	50.0	217	15	Q85j97 caprine art
	19	43	365	10	Q9m471 dendrobium
	20	43	370	10	Q9lw8 drosanthemu
	21	43	965	10	Q23946 gossypium h
	22	42	68	17	Q8tp58 methanosa
	23	42	339	10	Q85569 vitis vinif
	24	42	48.8	500	Q9exz2 mus musculu
	25	42	48.8	529	Q9qxz2 mus musculu
	26	42	48.8	530	Q8vcs0 mus musculu
	27	42	48.8	674	Q8s115 nicotiana s
	28	42	48.8	750	Q8s117 nicotiana s
	29	42	48.8	787	Q9x1w8 thermotoga
	30	42	48.8	960	Q9sa26 zea mays (m
	31	42	48.8	964	Q9cb3 lycopersico
	32	41	47.7	185	Q95sq9 drosophila
	33	41	47.7	185	Q9v3b7 drosophila
	34	41	47.7	712	Q8x0k3 neurospora
	35	41	47.7	716	Q53042 helicobacte
	36	41	47.7	757	Q49985 arabidopsis
	37	41	47.7	757	P92977 arabidopsis
	38	41	47.7	757	Q9c5m3 arabidopsis
	39	41	47.7	830	Q64551 arabidopsis
	40	41	47.7	971	Q967c6 haemonchus
	41	40.5	47.1	318	Q9rrv3 deinococcus
	42	40.5	47.1	469	Q51751 borrelia bu
	43	40.5	47.1	652	Q93346 caenorhabdi
	44	40	46.5	265	Q941d8 oryza sativ
	45	40	46.5	328	Q48623 triticum ae

ALIGNMENTS

RESULT 1					
Q9JLN4	ID	Q9JLN4	PRELIMINARY;	PRT;	183 AA.
	AC	Q9JLN4;			
	DT	01-OCT-2000 (T-EMBL1rel. 15, Created)			
	DT	01-OCT-2000 (T-EMBL1rel. 15, Last sequence update)			
	DT	01-OCT-2000 (T-EMBL1rel. 15, Last annotation update)			
	DE	Peptidoglycan recognition protein PGPR.			
	GN	PGPR.			
	OS	Rattus norvegicus (Rat).			
	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	OC	Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Rattus.			
	OX	NCBI_TAXID=10116;			
	RN	[1..]			
	RP	SEQUENCE FROM N.A.			
	RC	STRAIN=SPRAGUE-DAWLEY; TISSUE=SPLEEN;			
	RA	Rehman A., Teodorescu E.E., Krueger J.M.;			
	RL	Submitted (MAY 1999) to the EMBL/GenBank/DBJ databases.			
	DR	EMBL; AF154114; AAF73252.1; -			
	SO	SEQUENCE 183 AA; 20590 MW;			
		5B9C1B7AA8A2EC21 CRC64;			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	81	94.2	183	11	Q9JLN4	Q9JLN4 ratulus norv
2	73	84.9	193	6	Q9GK12	Q9GK12 camelus dro
3	72	83.7	190	6	Q8SPP7	Q8SPP7 bos taurus
4	50	58.1	148	11	Q9DAL7	Q9DAL7 mus musculu
5	49	57.0	190	5	Q9Vv97	Q9Vv97 drosophila
6	47	54.7	978	5	Q9USP6	Q9USP6 haemonchus
7	47	54.7	998	17	Q96552	Q96552 sulfolobus
8	46	53.5	215	17	Q8ZU01	Q8ZU01 pyrobaculum
9	46	53.5	432	2	Q93RK7	Q93RK7 bacillus sp
10	45	52.3	368	4	Q9HD75	Q9HD75 homo sapien
11	45	52.3	373	4	Q96LB8	Q96LB8 homo sapien
12	45	52.3	471	2	Q48204	Q48204 haemophilus
13	45	52.3	1478	5	Q9VB18	Q9vb18 drosophila
14	44.5	51.7	504	17	Q57984	Q57984 pyrococcus
15	44	51.2	341	4	Q96LB9	Q96LB9 homo sapien
16	44	51.2	460	2	Q48159	Q48159 haemophilus

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

score or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

OS	Camelus dromedarius (Dromedary) (Arabian camel).	GN	1700007L12Rik.
OC	Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OS	Mus musculus (Mouse).
Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
NCBI_TAXID=9838;	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.	
RN	[1]	OX	
RP	SEQUENCE FROM N.A.	RN	NCBI_TAXID=10090;
RC	TISSUE=LACTATING MAMMARY GLAND;	RP	[1]
RA	"Milk as a Source of Camel (Camelus dromedarius) Peptidoglycan Recognition Protein," Submitted (DEC 1998) to the EMBL/CenBank/DBJ/ DDBJ databases.	RC	SEQUENCE FROM N.A.
RT	"The Peptidoglycan recognition protein is expressed in the lactating mammary gland of camels and binds to lactic acid bacteria."	RC	STRAIN=CS7BL/6J; TISSUE=TESTIS;
RT	Submitted (MAR 2001) to the EMBL/CenBank/DBJ/ DDBJ databases.	RX	Medline=1217851;
RL	EMBL: AJ131676; CAC19553; 1; -.	RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Fukuda S., Adachi J., Kondo S., Saito T., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batyalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Garibaldi M., Gustincich S., Hill D., Hofmann M., Hume N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Suzuki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., RA
RN	EMBL: AJ409286; CAC84130; 1; -.	RA	"Functional annotation of a full-length mouse cDNA collection.";
RP	KW SIGNAL.	RT	Nature 408:685-690 (2000).
FT CHAIN	1 21 POTENTIN.	RL	DR
FT PEPTIDOGLYCAN RECOGNITION PROTEIN.	22 193 MW; B6A1BDB818030A7CB CRC64;	EMBL; AK005734; BAB2421.1; -.	
SEQUENCE	193 AA; 21377 MW;	DR	MGI; MGI:1922720; 1700007L12Rik.
SO	Query Match Score 73; DB 6; Length 193; Best Local Similarity 93.3%; Pred. No. 8.8e-05; Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	DR	InterPro; IPR000048; IQ_region.
Qy	2 GFLRSNYEYKGHRDV 16	SQ	RESULT 3
Db	157 GALRSNYEYKGHRDV 171	Q9VY97	SEQUENCE FROM N.A.
Q8SPP7	PRELIMINARY; PRT; 190 AA.	AC	09VY97; PRELIMINARY; PRT; 190 AA.
AC	08SPP7; PRT; 190 AA.	AC	09VY97; PRELIMINARY; PRT; 190 AA.
DT	01-JUN-2002 (TREMBLrel. 21, Created)	DT	01-MAY-2000 (TREMBLrel. 13, Created)
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)	DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE	01-JUN-2002 (TREMBLrel. 21, Last annotation update)	DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DT	Oligosaccharide-binding protein.	DE	CG9681 protein.
OS	Bos taurus (Bovine).	GN	PGRP-SB1 OR CG9681.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovidae; Bovidae; Bovinae; Bos.	OS	Drosophila melanogaster (Fruit fly).
OC	NCBI_TAXID=9913;	OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyrdoidea; Drosophilidae; Drosophila.
RN	[1]	RN	SEQUENCE FROM N.A.
RP	SEQUENCE FROM N.A.	Q9VY97	SEQUENCE FROM N.A.
RA	TydeLL C.C., Yount N.Y., Tran D., Yuan J., Selsted M.E.; "Isolation, characterization, and antimicrobial properties of bovine oligosaccharide-binding protein: A microbial granule protein of eosinophils and neutrophils."	AC	09VY97; PRELIMINARY; PRT; 190 AA.
RT	J. Biol. Chem. 0:0-(2002).	AC	09VY97; PRELIMINARY; PRT; 190 AA.
RT	DR	09DAL7	SEQUENCE FROM N.A.
RL	EMBL: AY083309; AAU87002; 1; -.	AC	SEQUENCE FROM N.A.
SQ	SEQUENCE 190 AA; 21063 MW;	AC	STRAIN=BERKELEY; PubMed=1073132;
Qy	1 RGFLRSNYEYKGHRDV 16	AC	Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scheer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazquez R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G., Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Balliew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
Db	152 RGFLRSNYEYKGHRDV 167	AC	NCBI_TAXID=7227;
Qy	1 RGFLRSNYEYKGHRDV 16	AC	NCBI_TAXID=BERKELEY;
Db	152 RGFLRSNYEYKGHRDV 167	AC	Medline=20196006;
Q9DAL7	PRELIMINARY; PRT; 148 AA.	AC	Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scheer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazquez R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G., Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Balliew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RESULT 4	09DAL7	AC	NCBI_TAXID=BERKELEY;
ID	01-JUN-2001 (TREMBLrel. 17, Created)	AC	Medline=20196006;
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)	AC	Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scheer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazquez R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G., Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Balliew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
DT	01-JUN-2002 (TREMBLrel. 17, Last sequence update)	AC	NCBI_TAXID=BERKELEY;
DT	1700007L12Rik protein.	AC	NCBI_TAXID=BERKELEY;

RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablo-B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Doudson K., Doupe L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W., Fosler C., Gabrilian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Hodges N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitus A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milishina N.V., Molarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzyk D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., Palazzolo M., Partman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reiner K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling C., Stapleton M., Strong R., Sun E., Svirsas R., Tecler C., Turner C., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C., "The genome sequence of <i>Drosophila melanogaster</i> ," Science 287:2185-2195(2000).	Qy	1 RGFLRSNYEVKGHRV 165	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
RA	AC 09U5P6; PRELIMINARY;	Qy	1 RGFLRSNYEVKGHRV 14	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
DB	152 RGYLKDNYTIFGHR 165	Db	152 RGYLKDNYTIFGHR 165	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
RESULT 6	Q9U5P6	Q9U5P6	Q9U5P6	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
ID	AC 09U5P6; PRELIMINARY;	ID	AC 09U5P6; PRELIMINARY;	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
AC	01-MAY-2000 (TREMBLrel. 1.3, Created)	AC	01-MAY-2000 (TREMBLrel. 1.3, Created)	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
DT	01-MAY-2000 (TREMBLrel. 1.3, Last sequence update)	DT	01-MAY-2000 (TREMBLrel. 1.3, Last sequence update)	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)	DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
DE	Microsomal aminopeptidases.	DE	Microsomal aminopeptidases.	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
GN	H11-L.	GN	H11-L.	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
OS	Haemonchus contortus (Barber pole worm).	OS	Haemonchus contortus (Barber pole worm).	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
OC	Eukaryota; Metazoa; Nematoidea; Chromadorea; Rhabditida; Strongylida; Haemonchidae; Haemonchinae; Haemonchus.	OC	Eukaryota; Metazoa; Nematoidea; Chromadorea; Rhabditida; Strongylida; Haemonchidae; Haemonchus.	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
DR	Submitted (OCT-1993) to the EMBL/GenBank/DDBJ databases.	DR	Submitted (OCT-1993) to the EMBL/GenBank/DDBJ databases.	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
EMBL	EMB24991; CAB57357; 1.	EMBL	EMB24991; CAB57357; 1.	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
MEROPS	M01.015;	MEROPS	M01.015;	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
InterPro	IPR001930; Ala_peptidase.	InterPro	IPR001930; Ala_peptidase.	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
PFAM	PF01433; Peptidase_M1.	PFAM	PF01433; Peptidase_M1.	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
PRINTS	PR00756; ALADPTASE.	PRINTS	PR00756; ALADPTASE.	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
PROSITE	PS00142; ZINC PROTEASE; UNKNOWN_1.	PROSITE	PS00142; ZINC PROTEASE; UNKNOWN_1.	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
KW	Aminopeptidase.	KW	Aminopeptidase.	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
SEQUENCE	978 AA;	SEQUENCE	978 AA;	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
SQ	112117 MW;	SQ	112117 MW;	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	978 AA.
RESULT 7	Q96Z52	Q96Z52	Q96Z52	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	998 AA.
ID	PRELIMINARY;	ID	PRELIMINARY;	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	998 AA.
AC	096Z52;	AC	096Z52;	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	998 AA.
DT	01-DEC-2001 (TREMBLrel. 19, Created)	DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	998 AA.
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)	DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	998 AA.
DE	Hypothetical protein STM1979.	DE	Hypothetical protein STM1979.	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	998 AA.
DN	SG	DN	SG	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	998 AA.
OS	Sulfolobaceae; Sulfolobales; Sulfolobaceae; Sulfolobus.	OS	Sulfolobaceae; Sulfolobales; Sulfolobaceae; Sulfolobus.	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	998 AA.
NCBI_TAXID	111955;	NCBI_TAXID	111955;	Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	998 AA.
SEQUENCE FROM N.A.		SEQUENCE FROM N.A.		Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	998 AA.
STRAIN=JCM 10545 / 7;		STRAIN=JCM 10545 / 7;		Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	998 AA.
PubMed=11572479;		PubMed=11572479;		Query Match	Score 49; DB 5; Length 190;	Best Local Similarity 57.1%; Pred. No. 1.4;	Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	PRT:	998 AA.
RA	Kawarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S., Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takanishi Y., Kato Y., Yoshihara T., Tanaka T., Kudo Y., Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A., Oshima T., Kikuchi H.; "Complete genome sequence of an aerobic thermoacidophilic	RA	Kawarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S., Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takanishi Y., Kato Y., Yoshihara T., Tanaka T., Kudo Y., Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A., Oshima T., Kikuchi H.; "Complete genome sequence of an aerobic thermoacidophilic	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
RA	RT	RT	RT	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
RA	REN	REN	REN	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
RA	RNA	RNA	RNA	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
RA	DNA	DNA	DNA	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
RA	DNA Res. 8:123-140(2001)	RA	DNA Res. 8:123-140(2001)	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
DR	AP00988; BAB6074.1;	DR	AP00988; BAB6074.1;	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
KW	Hypothetical protein Complete proteome.	KW	Hypothetical protein Complete proteome.	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
SQ	SEQUENCE FROM N.A.	SQ	SEQUENCE FROM N.A.	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
RA	Graham M., Smith T.S., Munn E.A., Coadwell W.J., Newton S.B., Barker P.J.J.; "Cloning and sequencing of variants of H11, the highly protective membrane glycoprotein from <i>Haemonchus contortus</i> ," Submitted (OCT-1993) to the EMBL/GenBank/DDBJ databases.	RA	Graham M., Smith T.S., Munn E.A., Coadwell W.J., Newton S.B., Barker P.J.J.; "Cloning and sequencing of variants of H11, the highly protective membrane glycoprotein from <i>Haemonchus contortus</i> ," Submitted (OCT-1993) to the EMBL/GenBank/DDBJ databases.	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
OC	Eukaryota; Metazoa; Nematoidea; Trichostomylida; Haemonchidae; Haemonchinae; Haemonchus.	OC	Eukaryota; Metazoa; Nematoidea; Trichostomylida; Haemonchidae; Haemonchinae; Haemonchus.	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
DR	DR	DR	DR	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
EMBL	EMBL	EMBL	EMBL	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
PROSITE	PROSITE	PROSITE	PROSITE	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
MEROPS	MEROPS	MEROPS	MEROPS	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
InterPro	IPR001930; Ala_peptidase.	InterPro	IPR001930; Ala_peptidase.	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
PFAM	PF01433; Peptidase_M1.	PFAM	PF01433; Peptidase_M1.	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
PRINTS	PR00756; ALADPTASE.	PRINTS	PR00756; ALADPTASE.	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
PROSITE	PS00142; ZINC PROTEASE; UNKNOWN_1.	PROSITE	PS00142; ZINC PROTEASE; UNKNOWN_1.	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
KW	Aminopeptidase.	KW	Aminopeptidase.	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
SEQUENCE	SEQUENCE	SEQUENCE	SEQUENCE	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
SQ	978 AA;	SQ	978 AA;	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
RESULT 8	Q8ZZXU1	Q8ZZXU1	Q8ZZXU1	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
ID	Q8ZZXU1; PRELIMINARY;	ID	Q8ZZXU1; PRELIMINARY;	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
AC	096Z52;	AC	096Z52;	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
DT	01-MAR-2002 (TREMBLrel. 20, Created)	DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)	DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
DE	Hypothetical protein PAE1101.	DE	Hypothetical protein PAE1101.	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
GN	Pyrobaculum aerophilum.	GN	Pyrobaculum aerophilum.	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
OC	Archaea; Crenarchaeota; Thermoproteales.	OC	Archaea; Crenarchaeota; Thermoproteales.	Query Match	Score 47; DB 17; Length 998;	Best Local Similarity 56.7%; Pred. No. 19;	Matches 10; Conservative 1; Mismatches 1; Indels 5; Indels 0; Gaps 0;	PRT:	998 AA.
DR	DR	DR</							

Query	Match	Score 46;	DB 17;	Length 215;
	Best Local Similarity	53.5%	No. 52;	
	Matches	60.0%	Pred.	
	Matches	9;	Mismatches	5;
	Conservative		Indels	0
2	GFLRSNVEVKGHRDV	16		
2Y				
87	GFLRLREKKGRREV	101		
DBB				

```

ID O96LB8; PRELIMINARY; PRT; 373 AA.
AC Q96LB8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Peptidoglycan recognition protein-I-beta precursor.
GN PGYLRPIBEA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBL_TAXID=9606;
RN [1] SEQUENCE FROM N.A.
RX MEDLINE=21433985; PubMed=11461926;
RA Liu C., Xu Z., Gupta D., Dziarski R. ;
RT "Peptidoglycan recognition proteins, a novel family of four human
innate immunity pattern recognition molecules." ;
RT J. Biol. Chem. 276:34686-4694 (2001).
RL EMBL; AY053377; AAK72485.1;
DR InterPro; IPR002086; Aldehyde_dehydr;
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 17 POTENTIAL.
SQ SEQUENCE 373 AA; 40662 MW; 14883166018A66AA CRC64;
Query Match 52.3%; Score 45; DB 4; Length 373;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 8; Conservative 3; Mismatches 5; Indels 0; G
QY 1 RGFIFRSNYEVKGHRDV 16
      :|:|| | :||| |
Db 336 KGYLTPNLYVGHSDV 351

RESULT 12
ID Q48204 PRELIMINARY; PRT; 471 AA.
AC Q48204;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE IgA1 protease (Fragment).
GN IgA.
OS Haemophilus influenzae.
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBL_TAXID=727;
RN [1] SEQUENCE FROM N.A.
RC STRAIN=HF284;
RX MEDLINE=95302961; PubMed=7783620;
RA Lomholt H., Poulsen K., Mogens K. ;
RT "Comparative characterization of the iga gene encoding IgA1 prote
in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilis
influenzae." ;
RT Mol. Microbiol. 15:495-506(1995).
RL X83487; CAA57870.1;
DR MEROPS; S06.001;
DR InterPro; IPR00710; IgA_S6.
DR Pfam; PF02395; IgA1; 1.
KW Procase.
FT NON_TER 1 1
FT NON_TER 471 471
SQ SEQUENCE 471 AA; 52559 MW; 68B57D31F21E6561 CRC64;

Query Match 52.3%; Score 45; DB 2; Length 471;
Best Local Similarity 61.5%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 4; Indels 0; G
QY 3 FLRSYEVKGHRD 15
      | :||| ||| |
Db 357 FFKGNYEVKGKTD 369

```


Tue Nov 5 11:08:12 2002

us-09-462-625-2_copy_145_160.rspt

Page 6

Query Match 51.2%; Score 44; DB 4; Length 341;
Best Local Similarity 53.3%; Pred. No. 19;
Matches 8; Conservative 2; Mismatches 5;
Matches 8;保守性 2; 错配 5;
Indels 0; Gaps 0;
QY 2 GFRLSNYEVKGRDY 16
Db 305 GYLTPNYLIMGHSDV 319

Search completed: November 4, 2002, 20:36:52
Job time : 97 secs

Percent Similarity:	93.33%	Conservative:	0
Best Local Similarity:	93.33%	Mismatches:	1
Query Match:	84.88%	Indels:	0
DB:	4	Gaps:	0
US-09-462-625-2_COPY_145_160 (1-16) x CD409286 (1-538)			
QY	2 GlypheLeuArgSerAsnTyrGluValLysGlyHisArgAspva 16		
Db	5151 GGAGCCCTGAGATCAAACGTCACAGGTCAAAGCACCGGGATGTG 5195		
RESULT 10			
LOCUS	AY083309 688 bp mRNA linear MAM 29-MAY-2002		
DEFINITION	Bos taurus oligosaccharide-binding protein mRNA, complete cds.		
REFERENCE	AY083309 1 (bases 1 to 688)		
AUTHORS	Tyndell,C.C., Yount,N., Tran,D., Yuan,J. and Selsted,M.E.		
TITLE	Isolation, characterization, and antimicrobial properties of bovine oligosaccharide-binding protein. A microbicidal granule protein of eosinophils and neutrophils		
JOURNAL	J. Biol. Chem. 277 (22), 19658-19664 (2002)		
PUBLINE	22280285		
REFERENCE	2 (bases 1 to 688)		
AUTHORS	Yount,N.J., Tyndell,C.C. and Selsted,M.E.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-MAR-2002) Pathology, UC Irvine, 440D Medical Sciences 1, Irvine, CA 92691, USA		
FEATURES			
source	Location/Qualifiers 1..688		
CDS	/organism="Bos taurus" /db_xref="taxon:9913" /feature_type="peripheral white blood cells" /translation="MSRRTPLAIVVLLGLGAAQDCGSIVSRKGWALASKCSORL RQPYYVVSHTAGSVNTPASCCQRAQNVQYYHREGRNCDVGNFLIGDLYVEG RGWNFLGAHSGPWNPPIAGISFNMHMVPASALRAQSLSLAGAARQYLTPYE VKGRDVOOTLSPDELYKTIQHRYRRV" /order(98..100..470..472) /note="microbicidal; undergoes N-terminal glutamine cyclization; peptidoglycan recognition protein; PGPR" /coor_start=1 /product="oligosaccharide-binding protein" /protein_id="BAI87002..1" /db_xref="GI:1955042"		
misc_feature	/note="disulfide bond" order(146..148..281..283)		
misc_feature	/note="disulfide bond" order(209..211..221..229)		
misc_feature	/note="disulfide bond" order(234..234..201..9..127..t)		
BASE COUNT	126 a 126 c 188 g 123 t		
ORIGIN			
Alignment Scores:			
red No.:	0.000271	Length:	688
Score:	72.00	Matches:	13
Percent Similarity:	87.50%	Conservative:	1
Best Local Similarity:	81.25%	Mismatches:	2
Query Match:	83.72%	Indels:	0
DB:	4	Gaps:	0
US-09-462-625-2_COPY_145_160 (1-16) x AF076483 (1-690)			
QY	1 ArgGlyPheLeuArgSerAsnTyrGluValLysGlyHisArgAspva 16		
Db	5116 CAGGAGCCGTGAGGTCAAATAGTGCAAAAGACACGGATGTG 563		
RESULT 12			
LOCUS	AF242517 724 bp mRNA linear SBB168 mRNA, complete cds.		
DEFINITION	Homo sapiens hypothetical protein AF242517		
VERSION	AF242517.1		
KEYWORDS			

SOURCE	Homo sapiens	FEATURES	information may be obtained at: http://www-bio.llnl.gov/bibp/genome/genome.html.
ORGANISM	Eukaryota; Malaria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarhini; Homidae; Homo.	source	Location/Qualifiers 1. .166500. /organism="Homo sapiens" /db_xref="taxon:9666" /chromosome="19" /map="19q13.3 between DM and C5R1." /clone="CIT-B-344H19" /clone="CIT-B-344H19" /note="L1NL clone name: BC282485" 70. .145 /rpt_family="TGAA"n complement(177..453 /rpt_family="AluJ"n complement(412..12) /rpt_family="MIR"n join(<746..904..1157..1239..4294..4451..5875..6036.. 1274..1294..13080..13162..13903..14053..14128, 14289..14568..14726..14804..14982..15039..15813..15885, 16025..16122) /note="Hypothetical arginine-rich gene product; Hypothetical CDS constructed from overlapping ESTs and Xgrail predictions. EST coverage is lacking for some portions of the CDS. Gene may be alternatively spliced, as multiple transcripts map into this region; alternatively, these individual clones may represent artificial events during RT and cDNA creation." /codon_start=3 /evidence="not experimental" /product="BC282485_1" /protein_id="AAD38244..1" /db_xref="GI:5042405" /translation="TLMQDPAQIQLQDYFRGVHAWRPAVGQVLELEVDRATDQ WRGIFDFADPTEIDTHKTGNFKOFNIFCHMLESALUTSSPSVTLDLTYTDLERSNRK MGGPGPSLAPRSQNLNSKEYLLILYSEEDRINPLPLYQEDPPVYLOGIILSKKE ELGRIGLGLDQNTTRDRENEIWHREQLASERLARLQLRQE AEALRVLGRGLRLERLRLERLQGRGQDCLRLAKELAERSLERSLRLKLTLT SELALKKRGRGSGSGRGRARPSPSPTGGRALRDTPTAVKAKERKORETOMKQORRN RLGSGSGGQPSWSRQFPAHLGGDAPNFSRQFSSVSDSFSSCASSSDL EDFESLSRGGRHRRKGKPSPTPWGSNMKSPPVERSHIQKSLANSGGWVPIKEYSE HOADMAMEIDARLKALQEYMNRLDMRS"
REFERENCE		/note="DOS similarity to overlapping ESTs: N39722 Yx92e04..r1 Homo sapiens cDNA clone 269214 5'; (12..110); 99% identity."-AA143147..214d07..r1 NCI_CGB_GCB1 Homo sapiens cDNA clone IMAGE:7131..01..5'; (6..163); 99% identity.-AA143160..214g067..r1 Soares pregnant uterus NBHPU Homo sapiens cDNA clone 505261..5'; (1..43); 100% identity.-AA14281..214g07..r1 Soares pregnant uterus NBHPU Homo sapiens cDNA clone 505261..3'; (570..527); 64% identity." 115'.1239 /note="predicted exon, program: grail2exons_human 1..3, frame: 0, quality: good, score: 62.000-DDS similarity to overlapping ESTs: N39722 Yx92e04..r1 Homo sapiens cDNA clone 269214 5'; (171..255); 95% identity."-AA283147..214d07..r1 NCI_CGB_GCB1 Homo sapiens cDNA clone IMAGE:7131..01..5'; (164..247); 92% identity.-AA143160..214g067..r1 Soares pregnant uterus NBHPU Homo sapiens cDNA clone 505261..5'; (44..126); 100% identity.-N28707 Yx66d1..r1 Homo sapiens cDNA clone 266709..5'; (1..66); 98% identity.-AA14281..214g067..r1 Soares pregnant uterus NBHPU Homo sapiens cDNA clone 505261..3'; (526..438); 82% identity." 1536..1634 /rpt_family="(CA)n" complement(1673..1973) /rpt_family="AluSx" complement(2040..2185) /rpt_family="LINE2" complement(2200..2498)	
AUTHORS	Lamerdin,J.B., McCreedy,P.M., Richardson,P., Sakaldasis,G., Burkhardt-Schultz,K., Gordon,L., Scott,D., Johnson,G., Stilwagen,S., Phan,H., Valasco,N., Do,L., Regala,W., Terry,A., Bangtan,L., Andrese,T., Amico-Keller,G., Coefield,A., Avila,J., Attix,C., Bruce,R., Thomas,P., Quan,G., Kronmiller,B., Arellano,A., Sanders,C., Ow,D., Nolan,M., Trong,S., Kobayashi,A., Olsen,A.S. and Carrano,A.V.	repeat_region	Direct Submission
JOURNAL	Submitted (11-JUN-1999) Joint Genome Institute, Lawrence Livermore National Laboratory, 7000 East Ave., Livermore, CA 94551, USA	repeat_region	Map and sequence oriented from q centromere to telomere of BAC 282485 overlaps cosmid R30477 to the left from bases 1 to 4,419 of this accession, and overlaps BAC 261576 (CIT-B-297N14; AC006262) from bases 146,203 to 166,500. Additional chromosome 19 map and sequence
COMMENT		repeat_region	
TITLE	Sequence analysis of a 2.5 Mb region in 19q13.2-13.3 between APOE and C5R1.	repeat_region	
JOURNAL	Unpublished	repeat_region	
REFERENCE	2 (bases 1 to 166500)	repeat_region	
AUTHORS	Lamerdin,J.B.	repeat_region	
JOURNAL	Direct Submission	repeat_region	
COMMENT	Submitted (11-JUN-1999) Joint Genome Institute, Lawrence Livermore National Laboratory, 7000 East Ave., Livermore, CA 94551, USA	repeat_region	
TITLE	Map and sequence oriented from q centromere to telomere of BAC 282485 overlaps cosmid R30477 to the left from bases 1 to 4,419 of this accession, and overlaps BAC 261576 (CIT-B-297N14; AC006262) from bases 146,203 to 166,500. Additional chromosome 19 map and sequence	repeat_region	

repeat_region /rpt_feature complement(2576..2746) repeat_region /rpt_feature complement(2747..3044) repeat_region /rpt_feature "AlusX" 3048..3750 repeat_region /rpt_feature "MER64" 3841..4060 repeat_region /rpt_feature "MER64" 4294..4451

/note= DDS similarity to overlapping ESTs: N39722 YX92e04.rl Homo sapiens cDNA clone 269214 5'; N39722 YX92e04.rl Homo sapiens cDNA clone 269214 5'; NCI-CGAP_GCB1 Homo sapiens cDNA clone IMAGE:713101 5'; (248..400); 99% identity.-AA143160. clone 505261 5'; (127..281); 97% identity.-N28707 YX66a11.rl Homo sapiens cDNA clone 2666709 5'; (67..222); 98% identity.-AA142881 z149d07.s1 Soares NbHPU Homo sapiens cDNA clone 505261 3'; (437..275); 97% identity."

complement(4789..5083)

repeat_region /rpt_feature "AlusX" 5875..5972

/note= DDS similarity to overlapping ESTs: N39722 YX92e04.rl Homo sapiens cDNA clone 269214 5'; (439..506); 71% identity.-AA143160 z149d07.rl Soares pregnant uterus NbHPU Homo sapiens cDNA clone 505261 5'; (282..379); 100% identity.-AA142881 z149d07.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 505261 3'; (274..375); 95% identity."

/note= DDS similarity to N28707.yx66d11.rl Homo sapiens cDNA clone 2666709 5'; (223..246); 100% identity.

5906..6036

/note= predicted exon, program: grail2exons_human_1.3, frame: 1, quality: excellent, score: 100.000"

repeat_region /rpt_feature complement(6148..6443)

repeat_region /rpt_feature complement(6608..6648)

repeat_region /rpt_feature "LINE2"

repeat_region /rpt_feature "MIR"

repeat_region /rpt_feature "(CA)n"

repeat_region /rpt_feature "MIR"

repeat_region /rpt_feature "AluJ0"

repeat_region /rpt_feature "MER5B"

repeat_region /rpt_feature "AlusX" 9909..9914

repeat_region /rpt_feature complement(10587..10752)

repeat_region /rpt_feature "AluJ"

repeat_region /rpt_feature "AlusX"

repeat_region /rpt_feature complement(11065..11191)

repeat_region /rpt_feature "AluJ" 11322..11619

repeat_region /rpt_feature "AlusX" 11775..11814

repeat_region /rpt_feature "AT_rich"

repeat_region /rpt_feature "AT_rich" complement(11864..11885)

repeat_region /rpt_feature "AT_rich" complement(11914..11976)

repeat_region /rpt_feature "Alus"

repeat_region /rpt_feature complement(12055..12088)

repeat_region /rpt_feature "AT_rich"

repeat_region complement(12186..12327)

repeat_region /rpt_feature complement(12334..12384)

repeat_region /rpt_feature "LINE2" 1284..12994

/note= predicted exon, program: grail2exons_human_1.3, frame: 1, quality: excellent, score: 92.000"

repeat_region /rpt_feature "GC_rich" 12847..12908

repeat_region /rpt_feature "GC_rich" 12895..12994

/note= DDS similarity to AA765166 nz79c05.s1 Homo sapiens cDNA clone IMAGE:1301672; (222..123); 97%

repeat_region /rpt_feature "GC_rich" 13002..13032

repeat_region /rpt_feature "GC_rich" 13080..13162

/note= predicted exon, program: grail2exons_human_1.3, frame: 2, quality: excellent, score: 100.000-DDS similarity to AA765166 nz79c05.s1 NCI-CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1301672; (122..48); 100%

identity."

repeat_region complement(13799..13954)

repeat_region /rpt_feature "GC_rich"

Alignment Scores:

Pred. No.:	3.64	Length:	166500
Score:	64.00	Matches:	12
Percent Similarity:	87.50%	Conservative:	2
Best Local Similarity:	75.00%	Mismatches:	2
Query Match:	74.42%	Indels:	0
DB:	9	Gaps:	0

US-09-462-625-2_COPY_145_160 (1-16) x AC007785 (1-166500)

QY 1 ArglyPheLeuArgSerAsnTyrgluValLysGlyHisArgAspVal 16

Db 17092 CAGGGAGCCCTGAGGTCCACTATGTCTCAAACGGACACCGGGATGTG 17045

RESULT 14

LOCUS	AC002319	46305 bp DNA	PRI 03-FEB-1998
DEFINITION	Homo sapiens chromosome 9q34,	clone 70C11,	complete sequence.
ACCESSION	AC002319		
VERSION	AC002319.1	GI:2828782	
KEYWORDS	HTG.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butherida; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 46305)		
AUTHORS	Birren, B., Fasman, K., McKernan, K., Nusbaum, C., Richardson, P. and Lander, E.		
TITLE	Homo sapiens chromosome 9q34, clone 70C11		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 46305)		
AUTHORS	Hawkins, T.L., Reeve, M.P., Christofferson, A., Birren, B.W., McKernan, K., Munro, C., Richardson, P., Barna, N., Chang, A., Cooke, P., Daly, M.J., Devon, K., Dewar, K., Forrester, C., Gage, D., Gerlengy, K., Guitau, G., Haug, B., Huang, J., Jacobson, L., Lane, M., Lee, K., MacKenzie, J., Marquis, N., McDermott, J., Molla, M., Moloney, N., Morrow, J., Nachman, A., Naylor, J., O'Connor, T., Olotu, A., Peterson, K., Rollins, G., Spender, J., Stillwell, J., Stone, C., Strickland, C., Sydney, K., Trainor, A., Wilmer, F., Zemtseva, I. and Zody, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (15-JUL-1997) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
REFERENCE	3 (bases 1 to 46305)		
AUTHORS	Birren, B., Fasman, K., McKernan, K., Nusbaum, C., Richardson, P., Lander, E., Allen, N., Baldwin, J., Baker, J., Byrne, S., Cantu, C., Castle, A., Cooke, P., Daly, M.J., Depaire, E., Devon, K., Dewar, K., Dutette, B., Etemadi, S., Ferreira, P., Forrest, C., Gage, D., Gardyna, S., Gensheimer, S.,		


```

repeat_region          rpt_family="TGAAAn"
repeat_region          complement(24150 . 24450)
repeat_region          rpt_family="AluSx" "24883"
repeat_region          complement(24854 . 24883)
repeat_region          rpt_family="AT_rich"
repeat_region          complement(24884 . 25173)
repeat_region          rpt_family="AluSx" "25199"
repeat_region          complement(25174 . 25199)
repeat_region          rpt_family="AT_rich"
repeat_region          complement(25385 . 25664)
repeat_region          rpt_family="AluO" "25921 . 26223
repeat_region          rpt_family="AluB" "26236 . 26522
repeat_region          rpt_family="AluSC" "27364 . 27497)
repeat_region          rpt_family="MIR" "26598 . 26903
repeat_region          complement(27024 . 27316)
repeat_region          rpt_family="AluSC" "27364 . 27497)

Alignment Scores:
Pred. No.:           114
Score:               53.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match:         DB: 9
Length:              46305
Matches:             7
Conservative:        7
Mismatched:          0
Indels:              0
Gaps:                0
US-09-462-625-2_COPY_145_160 (1-16) x AC002319 (1-46305)
QY      3 PheLeuArgSerAsnTyrGluValLysGlyHisArgAspVal 16
DB      : ::|||||:||||:||||:||||:||||:||||:||||:||||:||:|
RESULT 15
AL732364/C          AL732364 Human DNA sequence from clone RP11-430N14 on chromosome 9, complete
DEFINITION          sequence.
VERSION            AL732364.9
KEYWORDS           HTG
SOURCE             human.
ORGANISM           Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS            1 (bases 1 to 87903)
Lloyd,D
TITLE              Direct Submission
COMMENT           Submitted (17-AUG-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: hungerry@sanger.ac.uk clone requests: clonerequest@sanger.ac.uk On Aug 21, 2002 this sequence version replaced gi:22265441.
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: hungerry@sanger.ac.uk

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the

```

assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL, Sw:, SWISSPROT; Tr:, TRIMBL; Wp:, WORMPEP; information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr9 RP1-43N14 is from the library RPCI-11.2 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm VECTOR: pBAC3.6.
FEATURES
source          Location/Qualifiers
1. 1..87903 "Homo sapiens"
/db.xref="taxon:9606"
/chromosome="9"
/clone="RP11-430N14"
/clone.lib="RPCI-11.2"
BASE COUNT    23441 a 21236 c 22502 g 20724 t
ORIGIN
Alignment Scores:
Pred. No.:           230
Score:               53.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match:         DB: 9
Length:              230
Matches:             53
Conservative:        7
Mismatched:          0
Indels:              0
Gaps:                0
US-09-462-625-2_COPY_145_160 (1-16) x AL732364 (1-87903)
QY      3 PheLeuArgSerAsnTyrGluValLysGlyHisArgAspVal 16
DB      : ::|||||:||||:||||:||||:||||:||||:||||:||:|
Db      12761 TATTTAACAAACCATTCCTAAATCAAGGCATAGAGATTG 12720
Search completed: November 4, 2002, 19:54:28
Job time : 2804 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

SUMMARIES

OM protein - nucleic search, using frame_plus_p2n model	Run on: November 4, 2002, 17:49:26 ; Search time 301 Seconds	Result No.	Score	Query Match	Length	DB ID	Description
	{without alignments}	1	86	100.0	549	AAX21819	Mouse tag7 clone c
	119.708 Million cell updates/sec	2	86	100.0	677	18	Murine granulocyte
		3	72	83.7	688	18	Bovine granulocyte
Title: US-09-462-625-2_COPY_145_160		4	64	74.4	697	21	Human PRO1269 nucleic
Perfect score: 86		5	64	74.4	697	21	Human PRO1269 (UNQ)
Sequence: 1 RGFRLRSNEYVKGRHDV 16		6	64	74.4	697	22	DNA encoding prote
Scoring table: BLOSUM62		7	64	74.4	718	22	Human tag7 clone c
Xgapext 10.0 , Ygapext 0.5		8	64	74.4	726	21	Human Htg7 secret
Ygapop 10.0 , Ygapext 0.5		9	64	74.4	749	21	Chondrosarcoma Pep
Fgapop 6.0 , Fgapext 7.0		10	63	73.3	60	24	Human secreted pro
Delgap 6.0 , Delext 7.0		11	55	64.0	24	ABN38508	Human spliced tran
Searched: 2185239 seqs, 1125999159 residues		C 12	54	62.8	279	24	Human ORFX polyunc
Total number of hits satisfying chosen parameters: 4370478		C 13	49	57.0	23	AEI15529	Drosophila melanog
Minimum DB seq length: 0		C 14	49	57.0	2609	23	Drosophila melanog
Maximum DB seq length: 2000000000		C 15	49	57.0	2641	23	Drosophila melanog
Post-processing: Minimum Match 0%		C 16	48	55.8	65	24	Mouse spliced tran
Maximum Match 100%		C 17	48	55.8	621	22	C glutamicum codin
Listing first 45 summaries		C 18	48	55.8	744	22	Corynebacterium 91
		C 19	48	47.03	23	ABL02316	Drosophila melanog
		C 20	48	55.8	344980	22	C glutamicum codin
		C 21	47	54.7	716	22	Human ion channel-
		C 22	47	54.7	1689	14	AAH67897
		C 23	47	54.7	3006	14	AAH67897
		C 24	47	54.7	3084	14	AAH68533
		C 25	46	53.5	1513	23	AAH91071
		C 26	46	53.5	2263	22	Bacillus sp. KSM-K
		C 27	46	53.5	2880	22	Human immune/haema
		C 28	46	53.5	4480	18	Rat sarcoma virus
		C 29	46	53.5	4480	18	Nucleotide sequence
		C 30	46	53.5	18909	23	Drosophila melanog
		C 31	45.5	52.9	1429	22	S cerevisiae adopt
		C 32	45	52.3	392	22	Human colon cancer
		C 33	45	52.3	726	24	Human polynucleoti
		C 34	45	52.3	807	22	C glutamicum codin
		C 35	45	52.3	1110	22	Human full length
		C 36	45	52.3	1110	24	ABA90337
		C 37	45	52.3	1128	22	Human Polynucleoti
		C 38	45	52.3	1876	21	Human full length
		C 39	45	52.3	3395	24	Wound healing tiss
		C 40	45	52.3	4480	18	Listeria monocyt
		C 41	45	52.3	6913	23	Drosophila melanog
		C 42	45	52.3	349980	22	C glutamicum codin
		C 43	44	51.2	1092	23	DNA encoding novel
		C 44	44	51.2	7135	22	Tumour suppressor
		C 45	43	50.0	439	21	Zea mays DNA fragm
Database : N_Geneseq_10102,*							
1: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA1980.DAT:*							
2: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA1981.DAT:*							
3: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA1982.DAT:*							
4: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA1983.DAT:*							
5: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA1984.DAT:*							
6: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA1985.DAT:*							
7: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA1986.DAT:*							
8: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA1987.DAT:*							
9: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA1988.DAT:*							
10: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA1989.DAT:*							
11: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA1990.DAT:*							
12: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA1991.DAT:*							
13: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA1992.DAT:*							
14: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA1993.DAT:*							
15: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA1994.DAT:*							
16: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA1995.DAT:*							
17: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA1996.DAT:*							
18: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA1997.DAT:*							
19: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA1998.DAT:*							
20: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA2000.DAT:*							
21: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA2001A.DAT:*							
22: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA2001B.DAT:*							
23: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA2002.DAT:*							
24: /SIDS2/gcgdata/geneseq/geneseq -emb1/NA2002.sp.*							

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1	ID AAX21819	Standard: cDNA; 549 BP.
	XX	XX
	AC	AC
	XX	XX
	DT	18-MAY-1999 (first entry)
	XX	XX
	DE	Mouse tag7 clone coding sequence.
	XX	XX
	KW	KW
	KW	KW
	XX	melanoma; leukaemia; apoptosis inducer; mouse; ds.
	XX	XX
	OS	Mus sp.
	XX	XX
	PN	PN
	XX	WO9902686-A1.
	PD	21-JAN-1999.

FT CDS 29..601
 FT siq_peptide /tag= a
 FT mat_peptide /*tag= b
 FT mat_peptide 560..598
 FT /*tag= c
 FT /product= BGP-A
 XX WO20053750-A1.
 XX 21-AUG-1997.
 XX 97WO-US022218.
 XX 16-FEB-1996; 96US-0011834.
 XX (REGC) UNIV CALIFORNIA.
 XX Selsted ME;
 PI DR WPI; 1997-424753/39.
 DR P-PSDB; AAW3722.
 XX Antimicrobial (poly)peptide(s) from bovine and murine granulocytes -
 PT useful therapeutically, as preservatives for food, in water
 PT treatment and in agriculture
 XX
 PS Claim 13; Fig 4; 56pp; English.
 CC This cDNA sequence encodes the precursor (see AAW23722) of a novel,
 CC claimed antimicrobial peptide from bovine neutrophils, designated
 CC bovine granulocyte peptide A or BGP-A (see AAW3724). It was
 CC isolated from bovine bone marrow cDNA using PCR and RACE
 CC procedures. The encoded BGP-A precursor comprises a 21-amino acid
 CC signal peptide, a 156-residue propeptide and 13 C-terminal amino
 CC acids corresponding to mature BGP-A. BGP-A and the murine
 CC homologue, MGP-A (see AAW23725), exhibit activity against
 CC Gram-positive and Gram-negative bacteria, fungi and viruses,
 CC specifically *Staphylococcus aureus*, *Escherichia coli*, *Candida*
 CC *albicans*, *Salmonella typhimurium* and *C. neoformans* (claimed).
 CC They can be used in human or veterinary medicine (particularly to
 CC treat disorders associated with lipopolysaccharides e.g. sepsis
 CC and endotoxemia) or as preservatives in food products or in water
 CC supplies (claimed). They can also be applied to crops to reduce
 CC post-harvest spoilage or expressed in transgenic plants to increase
 CC their disease resistance. They have low immunogenicity.
 XX Sequence 688 BP; 126 A; 234 C; 201 G; 127 T; 0 other;
 SQ Alignment Scores:
 Pred. No.: 0.00028 Length: 688
 Score: 72.00 Matches: 13
 Percent Similarity: 87.50% Conservative: 1
 Best Local Similarity: 81.25% Mismatches: 2
 Query Match: 83.72% Indels: 0
 DB: 18 Gaps: 0
 US-09-462-625-2_COPY_145_160 (1-16) x AAC58104 (1-697)
 Qy 1 ArgGlyPheLeuArgSerAsnTyrGluValLysGlyHisArgAspVal 16
 DB 482 CGGGGATAACCTGACTCTAACTAGAACGGACCGCATGTG 529
 RESULT 4
 AAC58104 AAC58104 standard; cDNA; 697 BP.
 XX AAC58104;
 AC 25-JAN-2001 (first entry)
 XX Human PRO1269 nucleotide sequence SEQ ID NO:6.
 DE Human PRO1269 (UNQ639) cDNA sequence SEQ ID NO:215.
 KW Human; tumour; diagnosis; neoplastic disease; proliferation; cancer;
 KW identification; tumourigenesis; anticancer; detection; ss.
 XX OS Homo sapiens.
 XX PN WO20053750-A1.
 XX PD 14-SEP-2000.
 XX PF 02-DEC-1999; 99WO-US28851.
 XX PR 08-MAR-1999; 99WO-US05028.
 XX PR 01-SEP-1999; 99WO-US20111.
 XX PR 29-OCT-1999; 99US-0162506.
 XX PR 30-NOV-1999; 99WO-US28813.
 XX PR 01-DEC-1999; 99WO-US28634.
 PA (GETH) GENENTECH INC.
 PA Botstein D, Goddard A, Gurney AL, Roy MA, Watanabe CK, Wood WI;
 PI DR WPI; 2000-594320/56.
 DR P-PSDB; AAB24022.
 XX Antibodies specific for PRO polypeptides, used to diagnose and inhibit
 PT the growth of tumors in mammals, and to identify inhibitors of PRO
 PT polypeptide activity or expression -
 XX PS Claim 50; Fig 3; 226pp; English.
 XX The present invention describes an antibody that binds to a human
 CC protein (I) selected from: PRO381; PRO1269; PRO1410; PRO1755; PRO1780;
 CC PRO434; PRO1927; PRO3567; PRO1293; PRO1303; PRO434; PRO4354;
 CC PRO397; PRO4407; PRO1555; PRO1096; PRO2038; and PRO2262. (I) has
 CC antinecancer activity and can be used to diagnose tumors in mammals, by
 CC detecting complex formation when the antibody is contacted with test
 CC cells. Increased expression of genes encoding (I) can also be detected
 CC to diagnose tumors. Agents which inhibit the activity of (I),
 CC especially the antibodies, or an antisense oligonucleotide which
 CC hybridises to genes encoding (I), can be used to inhibit tumor growth,
 CC preferably by inducing cell death. Methods from the present invention
 CC can be used to identify compounds which inhibit the biological activity
 CC of (I). AAC58019 to AAC58102 represent PCR Primers and hybridisation
 CC probes used in examples from the present invention for human PRO
 CC sequences. AAC58103 to AAC58122 and AAC24021 to AAC24040 represent human
 CC PRO polynucleotide and protein sequences given in the exemplification of
 CC the present invention.
 XX Sequence 697 BP; 136 A; 250 C; 184 G; 127 T; 0 other;
 SQ Alignment Scores:
 Pred. No.: 0.00051 Length: 697
 Score: 64.00 Matches: 12
 Percent Similarity: 87.50% Conservative: 2
 Best Local Similarity: 87.50% Mismatches: 2
 Query Match: 74.42% Indels: 0
 DB: 21 Gaps: 0
 US-09-462-625-2_COPY_145_160 (1-16) x AAC58104 (1-697)
 Qy 1 ArgGlyPheLeuArgSerAsnTyrGluValLysGlyHisArgAspVal 16
 DB 497 CAGGGAGCCCTGAGTCCAACTAAGTCACGGATGTG 544
 RESULT 5
 AAA7082 ID AAA7082 standard; cDNA; 697 BP.
 XX AC AAA7082;
 DT 08-AUG-2000 (first entry)
 XX DE Human PRO1269 (UNQ639) cDNA sequence SEQ ID NO:215.

XX Human PRO polypeptide; membrane bound protein; receptor; diagnosis;
 KW transmembrane; secretion; immunoadhesion; pharmaceutical; screening;
 KW ss.

XX Homo sapiens.

XX WO200012708-A2.

XX PD 09-MAR-2000.

XX PR 01-SEP-1999; 99WO-US20111.

XX PR 01-SEP-1998; 98US-0098716.

XX PR 01-SEP-1998; 98US-0098749.

XX PR 02-SEP-1998; 98US-0098750.

XX PR 02-SEP-1998; 98US-0098833.

XX PR 02-SEP-1998; 98US-0098831.

XX PR 09-SEP-1998; 98US-0098834.

XX PR 09-SEP-1998; 98US-0099536.

XX PR 09-SEP-1998; 98US-0099536.

XX PR 09-SEP-1998; 98US-0099536.

XX PR 09-SEP-1998; 98US-0099632.

XX PR 10-SEP-1998; 98US-0099642.

XX PR 10-SEP-1998; 98US-0099741.

XX PR 10-SEP-1998; 98US-0099754.

XX PR 10-SEP-1998; 98US-0099763.

XX PR 10-SEP-1998; 98US-0099792.

XX PR 10-SEP-1998; 98US-0099838.

XX PR 10-SEP-1998; 98US-0099842.

XX PR 10-SEP-1998; 98US-0099845.

XX PR 10-SEP-1998; 98US-0099846.

XX PR 10-SEP-1998; 98US-0100285.

XX PR 15-SEP-1998; 98US-010338.

XX PR 15-SEP-1998; 98US-010390.

XX PR 16-SEP-1998; 98US-0100584.

XX PR 16-SEP-1998; 98US-0100637.

XX PR 16-SEP-1998; 98US-0100661.

XX PR 17-SEP-1998; 98US-0100661.

XX PR 17-SEP-1998; 98US-0100662.

XX PR 16-SEP-1998; 98US-0100664.

XX PR 17-SEP-1998; 98US-0100683.

XX PR 17-SEP-1998; 98US-0100584.

XX PR 17-SEP-1998; 98US-0100710.

XX PR 17-SEP-1998; 98US-0101011.

XX PR 18-SEP-1998; 98US-0101068.

XX PR 18-SEP-1998; 98US-0101071.

XX PR 22-SEP-1998; 98US-0101279.

XX PR 23-SEP-1998; 98US-0100930.

XX PR 18-SEP-1998; 98US-0100848.

XX PR 18-SEP-1998; 98US-0100849.

XX PR 18-SEP-1998; 98US-0101014.

XX PR 23-SEP-1998; 98US-0101475.

XX PR 23-SEP-1998; 98US-0101476.

XX PR 23-SEP-1998; 98US-0101477.

XX PR 24-SEP-1998; 98US-0101479.

XX PR 24-SEP-1998; 98US-0101730.

XX PR 29-SEP-1998; 98US-0102330.

XX PR 30-SEP-1998; 98US-0102331.

XX PR 30-SEP-1998; 98US-0102484.

XX PR 30-SEP-1998; 98US-0102487.

XX PR 30-SEP-1998; 98US-0102570.

XX PR 30-SEP-1998; 98US-0102571.

XX PR 01-OCT-1998; 98US-0102684.
 XX PR 01-OCT-1998; 98US-0102687.
 XX PR 02-OCT-1998; 98US-0102965.
 XX PR 06-OCT-1998; 98US-0103258.
 XX PR 06-OCT-1998; 98US-0103449.
 XX PR 07-OCT-1998; 98US-0103314.
 XX PR 07-OCT-1998; 98US-0103315.
 XX PR 07-OCT-1998; 98US-0103328.
 XX PR 07-OCT-1998; 98US-0103395.
 XX PR 07-OCT-1998; 98US-0103396.
 XX PR 07-OCT-1998; 98US-0103401.
 XX PR 08-OCT-1998; 98US-0103633.
 XX PR 08-OCT-1998; 98US-0103678.
 XX PR 08-OCT-1998; 98US-0103679.
 XX PR 14-OCT-1998; 98US-0104257.
 XX PR 20-OCT-1998; 98US-0104987.
 XX PR 20-OCT-1998; 98US-0105000.
 XX PR 21-OCT-1998; 98US-0105002.
 XX PR 21-OCT-1998; 98US-0105104.
 XX PR 22-OCT-1998; 98US-0105169.
 XX PR 22-OCT-1998; 98US-0105266.
 XX PR 26-OCT-1998; 98US-0105693.
 XX PR 26-OCT-1998; 98US-0105694.
 XX PR 27-OCT-1998; 98US-0105807.
 XX PR 27-OCT-1998; 98US-0105881.
 XX PR 27-OCT-1998; 98US-0105882.
 XX PR 27-OCT-1998; 98US-0106022.
 XX PR 28-OCT-1998; 98US-0106023.
 XX PR 28-OCT-1998; 98US-0106029.
 XX PR 28-OCT-1998; 98US-0106030.
 XX PR 28-OCT-1998; 98US-0106032.
 XX PR 28-OCT-1998; 98US-0106033.
 XX PR 28-OCT-1998; 98US-0106178.
 XX PR 29-OCT-1998; 98US-0106248.
 XX PR 29-OCT-1998; 98US-0106384.
 XX PR 29-OCT-1998; 98US-0106464.
 XX PR 30-OCT-1998; 98US-0106464.
 XX PR 03-NOV-1998; 98US-0106856.
 XX PR 03-NOV-1998; 98US-0106902.
 XX PR 03-NOV-1998; 98US-0106905.
 XX PR 03-NOV-1998; 98US-0106919.
 XX PR 03-NOV-1998; 98US-0106932.
 XX PR 03-NOV-1998; 98US-0106934.
 XX PR 10-NOV-1998; 98US-0107783.
 XX PR 17-NOV-1998; 98US-0108715.
 XX PR 17-NOV-1998; 98US-0108867.
 XX PR 17-NOV-1998; 98US-0108867.
 XX PR 17-NOV-1998; 98US-0108935.
 XX PR 18-NOV-1998; 98US-0108848.
 XX PR 18-NOV-1998; 98US-0108849.
 XX PR 18-NOV-1998; 98US-0108850.
 XX PR 18-NOV-1998; 98US-0108851.
 XX PR 18-NOV-1998; 98US-0108852.
 XX PR 18-NOV-1998; 98US-0108858.
 XX PR 18-NOV-1998; 98US-0108904.
 XX PA (GETH) GENENTECH INC.

XX XX Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;

XX XX DR WPI; 2000-237871/20.

XX XX P-PSDB; AAY9400.

XX PT New mammalian DNA sequences encoding transmembrane, receptor or secreted PRO polypeptides, useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interactions

XX
PS Claim 2; Fig 121; 773pp; English.
XX
CC AAI37022 to AAA37144 encode the new isolated human transmembrane receptor or secreted PRO polypeptides given in AA999340 to AAY99462. The transmembrane and receptor PRO proteins can be used for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interactions. The polypeptides and nucleotide sequences encoding them have various industrial applications, including uses as pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent PCR primers and hybridisation probes used in the isolation of the PRO polypeptides from the present invention.

XX Sequence 697 BP; 136 A; 250 C; 184 G; 127 T; 0 other;

XX Alignment Scores:

Pred. No. :	0.00951	Length:	697
Score:	64.00	Matches:	64.00
Percent. Similarity:	87.50%	Conservative:	87.50%
Best Local Similarity:	75.00%	Local Similarity:	75.00%
Query Match:	74.42%	Match:	74.42%
DB:	21	Mismatches:	2
		Indels:	0
		Gaps:	0

US-09-462-625-2_COPY_145_160 (1-16) x AAA37082 (1-697)

Qy 1 Arg Gly Phe Leu Arg Ser Asn Tyr Glu Val Lys Gly His Arg Asp Val 16

Db 497 CAGGGAGGCCCTGAGGTCCAACTATGTCCTCAAAGGGACACGGGATG 544

RESULT 6

AAF54356

ID AAF54356 standard; DNA; 697 BP.

AC AAF54356;

XX DT 02-APR-2001 (first entry)

XX DE DNA encoding protein of the invention #61.

XX KW Secreted; transmembrane; gene therapy; ss.

XX OS Unidentified.

XX PN WO200078961-A1.

XX PD 28-DEC-2000.

XX PF 18-FEB-2000; 2000WO-US04342.

XX PR 23-JUN-1999; 99US-0141037.

PR 20-JUL-1999; 99US-0141578.

PR 26-JUL-1999; 99US-0145698.

PR 01-SEP-1999; 99WO-US20111.

PR 29-OCT-1999; 99US-0161506.

PR 30-NOV-1999; 99WO-US28313.

PR 02-DEC-1999; 99WO-US28551.

PR 16-DEC-1999; 99WO-US30095.

PR 05-JAN-2000; 2000WO-US00219.

PR 06-JAN-2000; 2000WO-US00376.

XX PA (GETH) GENENTECH INC.

XX PI Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S, Gao W, Godowski PJ, Grimaldi CJ, Gurney AL, Hillian KJ, Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;

XX WPI; 2001-071395/08.

XX PT Secreted and transmembrane proteins and nucleic acids designated PRO useful as hybridization probes, in chromosome and gene mapping and gene

PT therapy -

XX PA

XX PS

XX

PS Claim 2; Fig 121; 787pp; English.

XX
CC The present invention relates to secreted and transmembrane proteins. These proteins and the DNA encoding them may be used as hybridization probes, in chromosome and gene mapping and in the generation of anti-sense RNA and DNA. They may also be used to generate either transgenic animals or knockout animals which are in turn useful for development and screening of therapeutically useful reagents. The nucleic acids may also be used in gene therapy.

XX Sequence 697 BP; 136 A; 250 C; 184 G; 127 T; 0 other;

XX Alignment Scores:

Pred. No. :	0.00951	Length:	697
Score:	64.00	Matches:	64.00
Percent. Similarity:	87.50%	Conservative:	87.50%
Best Local Similarity:	75.00%	Local Similarity:	75.00%
Query Match:	74.42%	Match:	74.42%
DB:	22	Mismatches:	2
		Indels:	0
		Gaps:	0

US-09-462-625-2_COPY_145_160 (1-16) x AAF54356 (1-697)

Qy 1 Arg Gly Phe Leu Arg Ser Asn Tyr Glu Val Lys Gly His Arg Asp Val 16

Db 497 CAGGGAGGCCCTGAGGTCCAACTATGTCCTCAAAGGGACACGGGATG 544

RESULT 7

AAX21820

ID AAX21820 standard; cDNA to mRNA; 718 BP.

XX AC AAX21820;

XX DT 18-MAY-1999 (first entry)

XX DE Human tag7 clone coding sequence.

XX KW Tag7; tumour growth inhibitor; mammalian tumour; carcinoma; sarcoma; melanoma; leukaemia; apoptosis inducer; human; ss.

XX OS Homo sapiens.

XX PN WO9902686-A1.

XX PD 21-JAN-1999.

XX PF 10-JUL-1998; 98WO-EP04287.

XX PR 11-JUL-1997; 97US-0893764.

XX PA (BOEHRINGER INGELHEIM INT GMBH.

XX PI Georgiev G, Ruselev S, Ostermann E, Prokhorchouk E;

XX DR WPI; 1999-120987/10.

XX DR P-PSDB; AAY00711.

XX PT New nucleic acid encoding tag7 - used to inhibit tumour growth and induce apoptosis, for treatment of carcinoma, sarcoma, melanoma and leukaemia.

XX PS Claim 11; Page 126-127; 138pp; English.

XX
CC This sequence encodes the human tag7 of the invention. Cells containing the tag7 DNA sequence are used to express recombinant tag7. Tag7 is used to produce and purify antibodies; to inhibit growth of mammalian tumours, especially for treating carcinoma (of liver, ovary, breast, cervix, lung, prostate, colon/rectum, bladder, testis, stomach, pancreas, mouth, head, and neck, squamous cell carcinoma or teratocarcinoma), sarcoma (Kaposi's, osteo- or fibro-sarcomas), melanoma or leukemia, and as a molecular weight marker. The tag7 polypeptide inhibits tumour growth and induces apoptosis. The tag7 coding sequences are also useful as probes for gene mapping and detection of tag7 gene expression, and as primers. Antibodies against tag7 are used as reagents for detecting tag7; as an antagonist of

CC tag7; for isolating tag7 and therapeutically to inhibit or delay tumour
 CC metastasis.
 XX Sequence 718 BP; 142 A; 251 C; 194 G; 131 T; 0 other;
 XX Alignment Scores:
 PRF. No.: 0.00985 Length: 718
 Score: 64.00 Matches: 12
 Percent Similarity: 87.50% Conservative: 2
 Best Local Similarity: 75.00% Mismatches: 2
 Query Match: 74.42% Indels: 0
 DB: 20 Gaps: 0
 US-09-462-625-2_COPY_145_160 (1-16) x AAX21820 (1-718)
 QY 1 ArgGlyPheLeuArgSerAsnTyrGluValIysGlyHisArgAspVal 16
 DB 524 CAGGGAGCCCTGAGGTCCAACATATGNGCTCAAAAGGACACCGGGATG 571
 RESULT 8
 ID AAA80613 Standard; cDNA; 726 BP.
 XX
 AC AAA80613;
 XX
 DT 21-NOV-2000 (first entry)
 XX Human Htag7 secreted protein gene #8.
 DE
 XX Secreted protein; immunosuppressive; anti-inflammatory; antiarthritic;
 KW anti-rheumatic; dermatological; antiproliferative; antiarteriosclerotic;
 KW anticancer; pulmonary; antiviral; antibacterial; antifungal;
 KW immune disorder; Addison's disease; rheumatoid arthritis; dermatitis;
 KW multiple sclerosis; inflammatory disorder; inflammatory bowel disease;
 KW Crohn's disease; nephritis; hyperproliferative disorder;
 KW cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer;
 KW melanoma; lymphoma; wound healing; human; ss.
 XX Homo sapiens.
 OS WO200029435-A1.
 PN
 PD 25-MAY-2000.
 XX
 PF 27-OCT-1999; 9WO-US25031.
 PR 28-OCT-1998; 98US-0105971.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ni J, Ruben SM, Olsen HS, Young PE, Kenny JJ, Moore PA, Wei Y;
 PI Greene JM;
 XX
 WPI: 2000-387742/33.
 DR P-PSDB; AAB25583.
 XX
 PT Isolated nucleic acid molecules encoding human secreted proteins are
 PT used for the prevention, amelioration and treatment of autoimmune,
 PT inflammatory, hyperproliferative and cardiovascular disorders, cancer,
 PT wounds, and infectious diseases -
 PS Claim 1: Figure 34: 803pp; English.
 XX
 CC The present invention relates to 12 secreted human proteins and the
 CC nucleotide sequences encoding them. The Polynucleotide sequences given
 CC in AAA80606-A80623 encode the 12 secreted protein sequences given in
 CC AAB25576-#25593. The human secreted proteins have various activities
 CC dependent on the tissues in which they are expressed. Examples of the
 CC activities of the proteins include: immunosuppressant;
 CC anti-inflammatory; antiarthritic; antineoplastic; dermatological;
 CC anti-proliferative; antiarteriosclerotic; anticancer; pulmonary;
 CC antiviral; antibacterial; and antifungal activity. The Proteins,
 CC polypeptides, agonists and antagonists may be used to treat prevent

CC and/or diagnose various disease, disorders and conditions examples of
 CC which include: immune disorders e.g. Addison's disease, rheumatoid
 CC arthritis, dermatitis, and multiple sclerosis; inflammatory disorders
 CC e.g. coronary arteriosclerosis and purpura;
 CC hyperproliferative disorders such as paroxysmaenia and myocarditis;
 CC cardiovascular disorders e.g. coronary arteriosclerosis and myocardiitis;
 CC cancer e.g. melanoma and lymphoma. The Proteins and Polynucleotide
 CC sequences may also be used in wound healing and the treatment of
 CC Infectious diseases. The human secreted protein gene #8 and protein
 CC sequences are represented in sequences AAA80613 and AAB25583. Sequences
 CC AAA8062-A80663 represent genes related to the secreted protein gene#8.
 XX SQ Sequence 726 BP; 161 A; 251 C; 188 G; 126 T; 0 other;
 XX Alignment Scores:
 PRD. No.: 0.00999 Length: 726
 Score: 64.00 Matches: 12
 Percent Similarity: 87.50% Conservative: 2
 Best Local Similarity: 75.00% Mismatches: 2
 Query Match: 74.42% Indels: 0
 DB: 21 Gaps: 0
 US-09-462-625-2_COPY_145_160 (1-16) x AAA80613 (1-726)
 QY 1 ArgGlyPheLeuArgSerAsnTyrGluValIysGlyHisArgAspVal 16
 DB 503 CAGGGAGCCCTGAGGTCCAACATATGNGCTCAAAAGGACACGGGATG 550
 RESULT 9
 AAA51719
 ID AAA51719 standard; cDNA; 749 BP.
 XX
 AC AAA51719;
 XX
 DT 31-OCT-2000 (first entry)
 DE Chondrosarcoma Peptidoglycan recognition protein-like protein cDNA.
 XX
 KW Peptidoglycan recognition protein-like protein; PGPR; PGPR-C; regulator;
 KW chondrosarcoma; immunosuppressive; antibacterial; anti-apoptotic;
 KW tumour necrosis factor; TNF; anti-infectious; mononuclear phagocyte;
 KW inhibitor; Protein co-ordinate data; ss.
 XX
 OS Homo sapiens.
 XX
 KEY FH Location/Qualifiers
 FT CDS 55..545
 FT /*tag= a
 FT /product= PGPR-like_protein
 FT 55..117
 FT /*tag= b
 FT mat_Peptide
 FT 118..642
 FT /*tag= c
 XX
 PN WO200039327-A1.
 XX
 PD 06-JUL-2000.
 XX
 PR 22-DEC-1999; 99WO-US30736.
 XX
 PR 23-DEC-1998; 98US-0113809.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Young PE, Olsen HS;
 XX
 DR WPI: 2000-452414/39.
 XX
 DR P-PSDB; AAY96964.
 XX
 PT Polynucleotide encoding peptidoglycan recognition protein-like protein,
 PT antibodies specific to it useful for preventing, treating conditions
 PT e.g. endotoxic shock and auto-immune disorders and infections in mammal
 XX

PS Claim 1; Fig 3; 191pp; English.

XX Novel human peptidoglycan recognition protein-like proteins (PGRP-W) expressed by keratinocytes (PGRP-K), wound healing tissues (PGRP-W) or chondrocytes tissue (PGRP-C) are disclosed. The PGPR DNA and proteins are useful for preventing, treating or ameliorating a medical system in a mammal (claimed). PGPR is useful in augmenting the immune system in such areas as immune recognition, antigen presentation and immune system activation. Antibodies or antagonists directed against these proteins may be useful in reducing or eliminating disorders associated with tumour necrosis factor (TNF) and TNF-like cytokines, such as endotoxic shock and autoimmune disorders and for treating infectious diseases including silicosis, sarcoidosis and idiopathic pulmonary fibrosis.

XX Sequence 749 BP; 165 A; 259 C; 195 G; 130 T; 0 other;

Alignment Scores:

Pred. No.:	0.0104	Length:	749
Score:	64.00	Matches:	12
Percent Similarity:	87.50%	Conservative:	2
Best Local Similarity:	75.00%	Mismatches:	2
Query Match:	74.42%	Indels:	0
DB:	21	Caps:	0

US-09-462-625-2_COPY_145_160 (1-16) x AAA51719 (1-749)

QY 1 ArgGlyPheLeuArgSerAsnTyrGluValLysGlyHisArgAspVal 16
Db :::||| |||||||:::||| |||||:::||| |||||:::||| |||||
526 CAGGGCCCTGAGGTCCAATATGTCTAAAGGRCACCGGGATGTG 573

RESULT 10
ID AAA80662 standard; cDNA; 285 BP.
XX
AC AAA80662;
XX DT 21-NOV-2000 (first entry)
XX DE Human secreted protein gene #8 related gene HBM79R SEQ ID #115
XX Secreted protein; immunosuppressant; anti-inflammatory; antiarthritic; anti-rheumatic; dermatological; antiprofenerative; anticancer; anticardiovascular; antihypertensive; antiviral; antibacterial; antifungal; immune disorder; Addison's disease; rheumatoid arthritis; dermatitis; multiple sclerosis; inflammatory disorder; inflammatory bowel disease; Crohn's disease; nephritis; hyperproliferative disorder; cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer; melanoma; lymphoma; wound healing; human; ss.
XX OS Homo sapiens.
XX PN WO200029435-A1.
XX PD 25-MAY-2000.
XX PF 27-OCT-1999; 99WO-US25031.
XX PR 28-OCT-1998; 98US-0105971.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Ni J, Ruben SM, Olsen HS, Young PE, Kenny JJ, Moore PA, Wei Y; Greene JM;
PI XX
DR 2000-387742/33.
XX PA Isolated nucleic acid molecules encoding human secreted proteins are used for the prevention, amelioration and treatment of autoimmune, inflammatory, hyperproliferative and cardiovascular disorders, cancer, wounds, and infectious diseases - Disclosure; Page 757; 803pp; English.

XX PS The present invention relates to 12 secreted human proteins and the nucleotide sequences encoding them. The polynucleotide sequences given in AAA80606-A80623 encode the 12 secreted protein sequences given in ARB5576-B25593. The human secreted proteins have various activities dependent on the tissues in which they are expressed. Examples of the activities of the proteins include: immunosuppressant; anti-inflammatory; antiarthritic; antihypertensive; antiviral; antibacterial; antifungal activity. The proteins, polypeptides, agonists and antagonists may be used to treat prevent and/or diagnose various disease, disorders and conditions examples of which include: immune disorders e.g. Addison's disease, rheumatoid arthritis, dermatitis, and multiple sclerosis; inflammatory disorders e.g. inflammatory bowel disease, Crohn's disease and nephritis; hyperproliferative disorders such as paraproteininaemias and purpura; cardiovascular disorders e.g. coronary arteriosclerosis and myocarditis; cancer e.g. melanoma and lymphoma. The proteins and polynucleotide sequences may also be used in wound healing and the treatment of infectious diseases. The human secreted protein gene # and protein sequences are represented in sequences AA80613 and AAB25583. Sequences AAA80662-A80663 represent genes related to the secreted protein gene#8.

SQ Sequence 285 BP; 71 A; 99 C; 62 G; 47 T; 6 other;

Alignment Scores:

Pred. No.:	0.00503	Length:	285
Score:	63.00	Matches:	12
Percent Similarity:	86.67%	Conservative:	1
Best Local Similarity:	80.00%	Mismatches:	2
Query Match:	73.26%	Indels:	0
DB:	21	Gaps:	0

US-09-462-625-2_COPY_145_160 (1-16) x AAA80662 (1-285)

QY 2 GlyPheLeuArgSerAsnTyrGluValLysGlyHisArgAspVal 16
Db 52 GGGGGCTGAGGTCCAATATGTCTAAAGGRCACCGGGATGTG 96

RESULT 11
ID ABN38508
ID ABN38508 standard; DNA; 60 BP.
XX AC ABN38508;
XX DT 15-JUL-2002 (first entry)
XX DE Human spliced transcript detection oligonucleotide SEQ ID NO:11256.
XX KW Human; mouse; rat; splice transcript; detection; RNA transcript; Homo sapiens.
XX KW splice variant; transcriptome; oligonucleotide library; ss.
XX KW Homo sapiens.
XX PN WO200210449-A2.
XX PD 07-FEB-2002.
XX PR 20-JUL-2001; 2001WO-1B01903.
XX PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX PA (COMP-) COMPUGEN INC.
XX PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX DR 2002-257383/30.
XX WPI 2002-257383/30.
XX PT New oligonucleotide libraries comprising oligonucleotides which selectively hybridize to mRNAs transcribed from a transcription unit of a genome, useful for detecting tissue-, pathology-, and developmental-specific genes -

XX Example 1; SEQ ID 11256; 47pp; English.
 XX
 CC The present invention describes oligonucleotide libraries for detecting messenger RNAs that populate a (sub-)transcriptome, where the (sub-)transcriptome comprises messenger RNAs transcribed from multiple transcription units that populate a genome. The library comprises several oligonucleotides, each capable of hybridising selectively to a set of messenger RNAs transcribed from a given transcription unit of the genome, which encodes one or more messenger RNA splice variants. The oligonucleotide libraries are useful for detecting mRNAs from a biological sample, in expression profiling studies, in qualitatively or quantitatively characterising the corresponding transcriptome, and in detecting RNA transcripts and splice variants of human or animal transcriptomes. The libraries may also be used as specialised mini-libraries to detect transcripts of a sub-transcriptome under a particular biological or pathological state, and so allowing the detection of tissue- and pathology-specific genes such as those genes only expressed in specific tissue under a specific pathological condition; to detect developmental specific genes; and to detect RNA transcripts and splice variants of a transcriptome of a patient suffering from a particular disorder. ABN2723, to ABN5989 represent Oligonucleotide sequences from rats, humans and mice, which are used in the exemplification of the present invention.
 CC N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences.
 XX Sequence 60 BP; 16 A; 18 C; 15 G; 11 T; 0 other;
 SQ

Alignment Scores:
 Pred. No.: 0.0258 Length: 60
 Score: 55.00 Matches: 10
 Percent Similarity: 91.67% Conservative: 1
 Best Local Similarity: 93.33% Mismatches: 1
 Query Match: 63.95% Indels: 0
 DB: 24 Gaps: 0
 US-09-462-625-2_COPY_145_160 (1-16) x ABN38508 (1-60)

QY 5 ArgSerAsnTyrGluValLysGlyHisArgAspVal 16
 Db 1 AGGTCCAACATAGTGTCAAAAGGACACGGGATGTG 36

RESULT 12

ID ABN24626/c
 XX ABN24626 standard; CDNA; 279 BP.
 AC ABN24626;
 DT 24-JUN-2002 (first entry)
 XX

DE Human ORFX polynucleotide sequence SEQ ID NO:17729.

XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis; hyperproliferative disorder; psoriasis; benign tumour; haemorrhage; degenerative disorder; osteoarthritis; neurodegenerative disorder; cardiovascular disease; hypothyroidism; cholesterol ester storage disease; immune deficiency; immune disorder; infectious disease; autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis; myasthenia gravis; gene; ss.
 OS Homo sapiens.
 XX WO200192533-A2.
 XX 06-DEC-2001.
 XX PF 29-MAY-2001; 2001WO-US10836.
 XX PR 30-MAY-2000; 2000US-206132P.
 PR 29-AUG-2000; 2000US-228716P.

XX (CURA-) CURAGEN CORP.
 XX PA
 XX P1 Shmikets RA, Leach MD;
 XX DR WPI: 2002-106308/14.
 XX DR P-PSDB; ABP08874.
 XX PT Novel human polypeptides and polynucleotides useful for diagnosing, preventing and treating cardiovascular disease, neurodegenerative, hyperproliferative disorders and autoimmune disorders -
 XX Disclosure: SEQ ID 17729; 1037BP; English.
 XX
 CC The present invention describes substantially purified human proteins (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1 in the specification). ABN15762 to ABN2723 encode the human ORFX proteins given in ABP0010 to ABP11500. ORFX proteins are useful for preventing or preventing pathology associated with an ORFX-associated disorder in humans, and in the manufacture of a medicament for treating a syndrome associated with ORFX-associated disorder. ORFX polynucleotide sequences can be used in gene therapy. ORFX sequences can be used in the treatment of cancer, hyperproliferative disorders, cirrhosis of liver, psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage, osteoarthritis, neurodegenerative disorders, disorders related to organ transplantation, cardiovascular diseases, diabetes mellitus, systemic lupus erythematosus, hypertension, hypothyroidism, cholesterol ester storage disease, various immune deficiencies and disorders, infectious diseases, autoimmune disorders such as multiple sclerosis, rheumatoid arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. ORFX proteins are also useful for treating burns, incisions, ulcers, for treating osteoporosis, bone degenerative disorders, or periodontal disease, and for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues and conditions resulting from systemic cytokine damage.
 CC N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 279 BP; 79 A; 73 C; 48 G; 79 T; 0 other;
 XX Alignment Scores:
 Pred. No.: 0.254 Length: 279
 Score: 54.00 Matches: 9
 Percent Similarity: 85.71% Conservative: 3
 Best Local Similarity: 64.29% Mismatches: 2
 Query Match: 62.79% Indels: 0
 DB: 24 Gaps: 0
 US-09-462-625-2_COPY_145_160 (1-16) x ABN24626 (1-279)
 QY 1 ArgGlyPheLeuArgSerAsnTyrgluValLysGlyHisArg 14
 DB 64 CGTGCGTTATCGGGCCAACTATGAATCAAAGAAACCGGA 23
 RESULT 13
 AB115529
 ID AB115529 standard; CDNA; 573 BP.
 AC AB115529;
 XX DT 26-MAR-2002 (first entry)
 XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 41069.
 XX KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX OS Drosophila melanogaster.
 XX PN WO200171042-A2.
 XX PR 06-DEC-2001.
 XX PF 29-MAY-2001; 2001WO-US10836.
 XX PR 30-MAY-2000; 2000US-206132P.
 PR 29-AUG-2000; 2000US-228716P.

PT Interactions -
 XX
 PS Claim 1; SEQ ID NO 41066; 21pp + Sequence Listing; English.
 PS
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 100 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB2072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 2641 BP: 785 A: 568 C: 598 G: 690 T: 0 other:
 Alignment Scores:
 Pred. No.: 33.9 Length: 2641
 Score: 49.00 Matches: 8
 Percent Similarity: 78.57% Conservative: 3
 Best Local Similarity: 57.14% Mismatches: 3
 Query Match: 56.98% Indels: 0
 DB: 23 Gaps: 0
 US-09-462-625-2-COPY_145_160 (1-16) x ABL15528 (1-2641)
 Qy 1 ArgGlyPheLeuArgSerAsnTyrGluValLysGlyIleArg 14
 |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
 Db 1522 CGGGATACCTCAAGGATAACTACACGCTGGTCGGTCATGG 1563

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 4, 2002, 19:00:46 ; Search time: 83 Seconds
(without alignments)
64,223 Million cell updates/sec

Title: US-09-462-625-2_COPY_145_160
Perfect score: 86
Sequence: 1 REFLRSNEYVRKGRDV 16

Scoring table: BLOSUM62
Xgapop 10.0 ' Xgapext 0.5
Ygapop 10.0 ' Ygapext 0.5
Fgapop 6.0 ' Fgapext 7.0
Delop 6.0 ' Delext 7.0

Searched: 310729 seqs, 166577418 residues
Total number of hits satisfying chosen parameters: 620558

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line Parameters:
-MODEL=frame_p2n.model -DEV=xlp
-Q=~cgn2.1/USPIO/spool/US09462625/runat_04112002.110249_23150/app/query.fasta_1.199
-DB=PUBLISHED_APPLICATIONS_NA -OPEN=faspar -SUFFIX=rnbp -MINMATCH=0.1
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdt -LIST=45 -DOCAUTL=200 -THR=MAX=100
-THR MIN=0 -ALIGN=L -MODE=L -LOCAL -OUTFILE=pt0 -NORM=ext -HEARSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USERID=US09462625 -CCGN=1_1_36_@runat_04112002.110249_23150
-NCPU=6 -ICPU=3 -NO_XLPYX -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT=LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -IGAPEXT=0.5 -DELCP=6 -DELEXT=7

Published Applications NA: *

- 1: /cgn2.6/podata/1/pubnpa/US07_PUBCOMB.seq:*
- 2: /cgn2.6/podata/1/pubnpa/PCT_NEW_PUB.seq:*
- 3: /cgn2.6/podata/1/pubnpa/US06_NEW_PUB.seq:*
- 4: /cgn2.6/podata/1/pubnpa/US05_PUBCOMB.seq:*
- 5: /cgn2.6/podata/1/pubnpa/US07_PUBCOMB.seq:*
- 6: /cgn2.6/podata/1/pubnpa/PCTUS_PUBCOMB.seq:*
- 7: /cgn2.6/podata/1/pubnpa/US08_NEW_PUB.seq:*
- 8: /cgn2.6/podata/1/pubnpa/US08_PUBCOMB.seq:*
- 9: /cgn2.6/podata/1/pubnpa/US09_NEW_PUB.seq:*
- 10: /cgn2.6/podata/1/pubnpa/US09_PUBCOMB.seq:*
- 11: /cgn2.6/podata/1/pubnpa/US10_NEW_PUB.seq:*
- 12: /cgn2.6/podata/1/pubnpa/US10_PUBCOMB.seq:*
- 13: /cgn2.6/podata/1/pubnpa/US10_PUBCOMB.seq:*
- 14: /cgn2.6/podata/1/pubnpa/US10_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Database :

Published Applications NA: *

```

5   43   50.0  17216  10  US-09-764-877-3565
6   43   50.0  17217  10  US-09-764-877-3566
7   42   48.8  307   10  US-09-780-253A-46
8   42   48.8  307   10  US-09-782-378A-26
9   41   47.7  143   10  US-09-983-955-2238
10  41   47.7  454   10  US-09-998-598-1444
11  41   47.7  472   10  US-09-970-444-152
12  41   47.7  1137  10  US-09-995-444-21
13  41   47.7  2081  10  US-09-771-16A-57
14  41   47.7  2211  10  US-09-812-7196
15  41   47.7  2211  10  US-09-815-242-7376
16  41   47.7  170334 10  US-09-835-232-7
17  40   46.5  185   10  US-09-978-574-122
18  40   46.5  256   10  US-09-5860-5860
19  40   46.5  817   10  US-09-966-881-36
20  40   46.5  1114  10  US-09-988-358-1
21  40   46.5  1546  10  US-09-988-358-2
22  40   46.5  1643  10  US-09-980-781-68
23  40   46.5  1758  10  US-09-994-128-88
24  40   46.5  2898  10  US-09-765-231A-41
25  40   46.5  2933  12  US-10-04-90-367
26  40   46.5  3050  10  US-09-925-310-350
27  39.5  45.9  17424 10  US-09-967-788A-314
28  39   45.3  273   10  US-09-960-322-1754
29  39   45.3  289   10  US-09-978-574-14787
30  39   45.3  473   10  US-09-917-800A-711
31  39   45.3  480   10  US-09-964-761-6550
32  39   45.3  964   10  US-09-974-300-1894
33  39   45.3  1192  10  US-09-980-107-1592
34  39   45.3  2036  10  US-09-729-674-79
35  39   45.3  3466  10  US-09-982-906-1
36  39   45.3  4389  12  US-10-016-388-1
37  39   45.3  143306 10  US-09-729-90-3
38  39   45.3  39758  10  US-09-913-300-3
39  38.5  44.8  181   10  US-09-963-817-461
40  38   44.2  180   10  US-09-964-761-1835
41  38   44.2  207   10  US-09-974-300-5121
42  38   44.2  233   10  US-09-960-352-144891
43  38   44.2  234   10  US-09-967-701-8474
44  38   44.2  235   10  US-09-978-574-510
45   44.2  385   10  US-09-878-574-1471
```

ALIGNMENTS

```

US-09-263-959-1/C
; Sequence 1, Application US/09263959
; Patent No. US2002105089A1
; GENERAL INFORMATION:
;   APPLICANT: Hood, Leroy E.
;   APPLICANT: Rowan, Lee
;   APPLICANT: Koop, Ben F.
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH
; NUMBER OF SEQUENCES: 12/9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; APPLICATION NUMBER: US/09/263,959
; FILING DATE: 05-MAR-1999
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
```

```

; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 920010.426C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4300
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6849/3 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-263-359-1

RESULT 2
US-09-462-625-2_COPY_145_160 (1-16) x US-09-263-959-1 (1-684973)
; Sequence 14908, Application US/09864761.
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanelz, David K.
; APPLICANT: Chen, Wenshang
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864-761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21

```

```

; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Amnonax Sequence Listing Engine vers. 1.1
; SEQ ID NO: 14908
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004832.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
US-09-864-761-14908

Alignment Scores:
Pred. No.: 4.04e+04 Length: 684973
Score: 44.00 Matches: 8
Percent Similarity: 69.23% Conservative: 1
Best Local Similarity: 61.54% Mismatches: 4
Query Match: 51.16% Indels: 0
DB: 10 Gaps: 0

US-09-462-625-2_COPY_145_160 (1-16) x US-09-864-761-14908 (1-550)

Qy 3 PheLeuArgSerAsnTygGluValLysIleGlyHisArgAsp 15
Db 502 TACCTAGGAGTCGGCCATCGTGT 540

RESULT 3
US-09-822-830A-424
; Sequence 424, Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
; APPLICANT: Genetics Institute, Inc.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Rechtein, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Guilukota, Kamalakar
; APPLICANT: Graham, James R.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6402
; CURRENT APPLICATION NUMBER: US/09/822-830A
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,604
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SEQ ID NO: 424
; LENGTH: 2228
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2055,2076,2140,2171,2172,2174,2190
; OTHER INFORMATION: n=a,c,g, or t
US-09-822-830A-424

Alignment Scores:
Pred. No.: 63.5 Length: 2228
Score: 43.00 Matches: 7
Percent Similarity: 83.33% Conservative: 3
Best Local Similarity: 58.33% Mismatches: 2
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0

US-09-462-625-2_COPY_145_160 (1-16) x US-09-822-830A-424 (1-2228)

```

Qy 5 ArgSerAsnTyrGluValLysGlyHisArgAspVal 16
 Db 1640 AGCACACATACTGTGAAAGGCCACAGGACCTC 1675

RESULT 4
 US-10-054-691-1/c

; Sequence 1, Application US/10054691
 ; Patent No. US2002011546A1

; GENERAL INFORMATION:

; APPLICANT: Yu, Xuanchuan

; ATTORNEY: Miranda, Maricar

; TITLE OF INVENTION: No. US2002011546A1 Human Lipase and Polynucleotides Encoding t

; FILE REFERENCE: LEY-003-US-A

; CURRENT FILING DATE: 2001-01-22

; PRIOR APPLICATION NUMBER: US 60/264,049

; PRIOR FILING DATE: 2001-01-24

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 4377

; TYPE: DNA

; ORGANISM: homo sapiens

; US-10-054-691-1

Alignment Scores:

Pred. No.: 144 Length: 4377
 Score: 43.00 Matches: 7

Percent Similarity: 76.32% Conservative: 3

Best Local Similarity: 53.85% Mismatches: 3

Query Match: 50.00% Indels: 0

DB: 12 Gaps: 0

US-09-462-625-2_COPY_15_160 (1-16) x US-10-054-691-1 (1-4377)

Qy 4 LeuArgSerAsnTyrGluValLysGlyHisArgAspVal 16
 Db 1769 TTATCATCAAATCTAGGACACAGGACAGAACCTG 1731

RESULT 5
 US-09-764-877-3565

; Sequence 3565, Application US/09764877

; Patent No. US20020147140A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC005

; CURRENT APPLICATION NUMBER: US/09/764,877

; CURRENT FILING DATE: 2001-01-17

; PRIOR application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 4031

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 17216

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-09-764-877-3565

Alignment Scores:

Pred. No.: 751 Length: 17216
 Score: 43.00 Matches: 8

Percent Similarity: 76.92% Conservative: 2

Best Local Similarity: 61.54% Mismatches: 3

Query Match: 50.00% Indels: 0

DB: 10 Gaps: 0

US-09-462-625-2_COPY_145_160 (1-16) x US-09-764-877-3565 (1-17216)

Qy 3 PheLeuArgSerAsnTyrGluValLysGlyHisArgAspVal 15
 Db 15183 TACCTAAGGAGTCCCGAGGACTGTCGTCATCGTGTAT 152221

RESULT 6
 US-09-764-877-3566
 ; Sequence 3566, Application US/09764877
 ; Patent No. US20020147140A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; FILE REFERENCE: PC005
 ; CURRENT APPLICATION NUMBER: US/09/764,877
 ; CURRENT FILING DATE: 2001-01-17
 ; PRIOR application data removed - refer to PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 4031
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 17217
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-764-877-3566
 Alignment Scores:
 Pred. No.: 752 Length: 17217
 Score: 43.00 Matches: 8
 Percent Similarity: 76.92% Conservative: 2
 Best Local Similarity: 61.54% Mismatches: 3
 Query Match: 50.00% Indels: 0
 DB: 10 Gaps: 0
 US-09-462-625-2_COPY_145_160 (1-16) x US-09-764-877-3566 (1-17217)
 Qy 3 PheLeuArgSerAsnTyrGluValLysGlyHisArgAspVal 15
 Db 15183 TACCTAAGGAGTCCCGAGGACTGTCGTCATCGTGTAT 152221
 RESULT 7
 US-09-880-253A-46/c
 ; Sequence 46, Application US/09880253A
 ; Patent No. US200201132A1
 ; GENERAL INFORMATION:
 ; APPLICANT: The University of Queensland
 ; TITLE OF INVENTION: EXPRESSION MODULATING SEQUENCES
 ; FILE REFERENCE: 2415281/EUH
 ; CURRENT APPLICATION NUMBER: US/09/880,253A
 ; CURRENT FILING DATE: 2001-06-13
 ; NUMBER OF SEQ ID NOS: 60
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 46
 ; LENGTH: 307
 ; TYPE: RNA
 ; ORGANISM: mouse
 ; US-09-880-253A-46
 Alignment Scores:
 Pred. No.: 8.94 Length: 307
 Score: 4.200 Matches: 7
 Percent Similarity: 75.00% Conservative: 5
 Best Local Similarity: 43.75% Mismatches: 4
 Query Match: 48.84% Indels: 0
 DB: 10 Gaps: 0
 US-09-462-625-2_COPY_145_160 (1-16) x US-09-880-253A-46 (1-307)
 Qy 1 ArgGlyPheLeuArgSerAsnTyrGluValLysGlyHisArgAspVal 16
 Db 93 AAGGGAGAGATTCATGATGTTGGAGAGCGGCAGATAGAGACATC 46
 RESULT 8
 US-09-782-378A-26/c
 ; Sequence 26, Application US/09782378A
 ; Patent No. US20020102731A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hearing, Patrick
 ; APPLICANT: Bahou, Wade
 ; APPLICANT: Sandalon, Ziv

APPLICANT: Gnatenko, Dmitri
 TITLE OF INVENTION: Adenoviral Vectors
 FILE REFERENCE: STONYB-04970
 CURRENT FILING DATE: 2001-09-09/782,378A
 PRIORITY APPLICATION NUMBER: 60/237,747
 PRIOR FILING DATE: 2000-10-02
 NUMBER OF SEQ ID NOS: 27
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 26
 LENGTH: 35100
 TYPE: DNA
 FEATURE: Human adenovirus type 17
 NAME/KEY: misc_feature
 LOCATION: (25184)..(25184)
 OTHER INFORMATION: The n at this position can be a, c, t, or g.
 US-09-782-378A-26

Alignment Scores:
 Pred. No.: 2.75e+03 Length: 35100
 Score: 42.00 Matches: 8
 Percent Similarity: 71.43% Conservative: 2
 Best Local Similarity: 57.14% Mismatches: 4
 Query Match: 48.84% Indels: 0
 DB: 10 Gaps: 0

US-09-462-625-2_COPY_145_160 (1-16) x US-09-782-378A-26 (1-35100)

Qy 1 ArgGlyPheLeuArgSerAsnTyrgluValLysGlyHisArg 14
 Db 8971 CGTGATTCATCGGGCCATGGAAAGGCTGGACCATCG 8930

RESULT 9
 US-09-983-965-2238 Sequence 2238, Application US/09983965
 PRIORITY APPLICATION NUMBER: US20030137160A1
 GENERAL INFORMATION:
 APPLICANT: Warren, Wesley C.
 APPLICANT: Tao, Ningbing
 APPLICANT: Byatt, John C.
 APPLICANT: Mathialagan, Nagappan
 TITLE OF INVENTION: NUCLEAR ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
 TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
 FILE REFERENCE: 37-21(10297)C
 CURRENT APPLICATION NUMBER: US/09/983,965
 CURRENT FILING DATE: 2001-10-26
 PRIOR APPLICATION NUMBER: US/09/465,231
 PRIOR FILING DATE: 1999-12-15
 PRIOR APPLICATION NUMBER: US 60/113,678
 PRIOR FILING DATE: 1998-12-17
 NUMBER OF SEQ ID NOS: 5912
 SEQ ID NO: 238
 LENGTH: 143
 TYPE: DNA
 ORGANISM: Bos taurus
 FEATURE:
 OTHER INFORMATION: Clone ID: 50-LIB3057-002-Q1-K1-E6
 US-09-983-965-2238

Alignment Scores:
 Pred. No.: 5.49 Length: 143
 Score: 41.00 Matches: 7
 Percent Similarity: 90.00% Conservative: 2
 Best Local Similarity: 70.00% Mismatches: 1
 Query Match: 47.67% Indels: 0
 DB: 10 Gaps: 0

US-09-462-625-2_COPY_145_160 (1-16) x US-09-983-965-2238 (1-143)

Qy 2 GlyPheLeuArgSerAsnTyrgluValLys 11
 Db 73 GGCTTATAAGATCTAACATCATTTAA 102

RESULT 10
 US-09-998-1440 Sequence 1440, Application US/09998598
 PATENT NO. US20020150922A1
 GENERAL INFORMATION:
 APPLICANT: Stolk, John A.
 APPLICANT: Xu, Jiangchun
 APPLICANT: Chenault, Ruth A.
 APPLICANT: Meagher, Madeline Joy
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
 FILE REFERENCE: 210121..561
 CURRENT APPLICATION NUMBER: US/09/998,598
 CURRENT FILING DATE: 2001-11-16
 NUMBER OF SEQ ID NOS: 2606
 SOFTWARE: Corixa Invention Disclosure Database
 SEQ ID NO: 1440
 LENGTH: 454
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-998-1440

Alignment Scores:
 Pred. No.: 22.2 Length: 454
 Score: 41.00 Matches: 7
 Percent Similarity: 71.43% Conservative: 3
 Best Local Similarity: 50.00% Mismatches: 4
 Query Match: 47.67% Indels: 0
 DB: 10 Gaps: 0

US-09-462-625-2_COPY_145_160 (1-16) x US-09-998-1440 (1-454)

Qy 1 ArgGlyPheLeuArgSerAsnTyrgluValLysGlyHisArg 14
 Db 102 CGTCCTTTCATCAATTACACTGTATGGTATAAAAGTCACCGT 143

RESULT 11
 US-09-770-444-152 Sequence 152, Application US/0977044
 PATENT NO. US2002023280A1
 GENERAL INFORMATION:
 APPLICANT: Gorlach, Jorn
 APPLICANT: An, Yong-Olang
 APPLICANT: Hamilton, Carol M.
 APPLICANT: Price, Jennifer L.
 APPLICANT: Raines, Tracy M.
 APPLICANT: Yu, Yang
 APPLICANT: Rameaka, Joshua G.
 APPLICANT: Page, Amy
 APPLICANT: Matthew, Abraham V.
 APPLICANT: Ledford, Brooke L.
 APPLICANT: Woessner, Jeffrey P.
 APPLICANT: Haas, William David
 APPLICANT: Garcia, Carlos A.
 APPLICANT: Ricker, Maja
 APPLICANT: Slader, Ted
 APPLICANT: Davis, Keith R.
 APPLICANT: Allen, Keith
 APPLICANT: Hoffman, Neil
 APPLICANT: Hurban, Patrick
 TITLE OF INVENTION: Expressed Sequences of Arabidopsis
 FILE REFERENCE: 2027 (PARA-016PRV)
 CURRENT APPLICATION NUMBER: US/09/770,444
 CURRENT FILING DATE: 2001-01-26
 PRIOR APPLICATION NUMBER: 80/178,502
 PRIOR FILING DATE: 2000-01-27
 NUMBER OF SEQ ID NOS: 999
 SOFTWARE: PastSeq for Windows Version 4.0
 SEQ ID NO: 152
 LENGTH: 472
 TYPE: DNA

```

; - ORGANISM: Arabidopsis thaliana
US-09-770-444-152

Alignment Scores:
Pred. No.: 23 .2 Length: 472
Score: 41.00 Matches: 8
Percent Similarity: 68.75% Conservative: 3
Best Local Similarity: 50.008 Mismatches: 5
Query Match: 47.67% Indels: 0
DB: 10 Gaps: 0

RESULT 1.2
US-09-905-494-21
Sequence 21, Application US/0995494
; GENERAL INFORMATION:
; Patent No. US2002012758A1
; PRIORITY NUMBER: US/0995494
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
FILE REFERENCE: TX-093
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 60/253,176
; PRIOR FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: Patentin version 3.1
; SEQ ID NO: 21
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-995-494-21

Alignment Scores:
Pred. No.: 23 .2 Length: 1137
Score: 41.00 Matches: 7
Percent Similarity: 88.89% Conservative: 1
Best Local Similarity: 77.78% Mismatches: 1
Query Match: 47.67% Indels: 0
DB: 10 Gaps: 0

RESULT 1.3
US-09-771-161A-57
Sequence 57, Application US/09771161A
; GENERAL INFORMATION:
; Patent No. US2002010811A1
; PRIORITY NUMBER: US/09771161A
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 80:620-2005.1
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/774,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: Patentin version 3.0
; SEQ ID NO: 57
; LENGTH: 2081
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-771-161A-57

Alignment Scores:
Pred. No.: 140 Length: 2081
Score: 41.00 Matches: 7
Percent Similarity: 75.00% Conservative: 2
Best Local Similarity: 58.33% Mismatches: 3
Query Match: 47.67% Indels: 0
DB: 10 Gaps: 0

RESULT 1.4
US-09-815-242-7196
Sequence 7196, Application US/09815242
; GENERAL INFORMATION:
; Patent No. US2002006156A1
; PRIORITY NUMBER: US/09815242
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; PROKARYOTES
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/194,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 7196
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE: CDS
; NAME/KEY: CDS
; LOCATION: (1)...(2211)
; US-09-815-242-7196

Alignment Scores:
Pred. No.: 150 Length: 2211
Score: 41.00 Matches: 11
Percent Similarity: 65.00% Conservative: 2
Best Local Similarity: 55.00% Mismatches: 3
Query Match: 47.67% Indels: 4
DB: 10 Gaps: 2

```

US-09-462-625-2_COPY_145_160 (1-16) x US-09-815-242-7196 (1-2211)

Qy 1 ArgGlyPheLeuArgSerAsnTyrGluVal-----LysGlyHis---ArgAspVal 16
 Db 49 AAAATTTGGATAAAATATGAAGTCATGGCTCTAAAGGCATGTAGGGATTAA 108

RESULT 15

US-09-815-242-7376

Sequence 7376, Application US/09815242

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Wall, Daniel

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; TITLE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.OLIA

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS.: 14110

; SOFTWARE: FastSEQ For Windows Version 4.0

; SEQ ID NO: 7376

; LENGTH: 2211

; TYPE: DNA

; FEATURE: ORGANISM: Helicobacter pylori

; NAME/KEY: CDS

; LOCATION: (1)...(2211)

US-09-815-242-7376

; Alignment Scores:

Pred. No.:	150	Length:	2211
Score:	41.00	Matches:	11
Percent Similarity:	65.00%	Conservative:	2
Best Local Similarity:	55.00%	Mismatches:	3
Query Match:	47.67%	Indels:	4
DB:	10	Gaps:	2

US-09-462-625-2_COPY_145_160 (1-16) x US-09-815-242-7376 (1-2211)

Qy 1 ArgGlyPheLeuArgSerAsnTyrGluVal-----LysGlyHis---ArgAspVal 16
 Db 49 AAAATTTGGATAAAATATGAAGTCATGGCTCTAAAGGCATGTAGGGATTAA 108

Search completed: November 4, 2002, 20:33:21
 Job time : 128 secs

Alignment Scores:																	
Pred. No.:		5.53e-07		Length:		669											
Score:		86.00		Matches:		16											
Percent Similarity:		100.00%		Conservative:		0											
Best Local Similarity:		100.00%		MisMatches:		0											
Query Match:		100.00%		Indels:		0											
DB:		10		Gaps:		0											
REFERENCE	US-09-462-625-2_COPY_145_160 (1-16) x AF193843 (1-669)																
AUTHORS	Georgiev, G.P., Kiselev, S.L., Prokhorchouk, E.B. and Ostermann, E.																
TITLE	Nucleic acid encoding tag7 polypeptide																
JOURNAL	Patent: US 6172211-A 109 JAN-2001;																
FEATURES	Location/Qualifiers																
SOURCE	1..549 /organism="unknown"																
BASE COUNT	117 a 161 c 151 g 120 t																
ORIGIN	/organism="Mus musculus"																
RESULT 2	MMRNATMS1																
LOCUS	MMRNATMS1																
DEFINITION	mRNATMS1																
ACCESSION	M.musculus mRNA for TAG7 protein.																
VERSION	X86374																
KEYWORDS	X86374..1 GI:887524																
SOURCE	tms1 gene.																
ORGANISM	Mus musculus.																
MATERIAL	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.																
REFERENCE	1 (base 1 to 678)																
AUTHORS	Kustikova, O.S., Kiselev, S.L., Borodulina, O.R., Senin, V.M., Afanaseva, A.V. and Kabishev, A.A.																
TITLE	Cloning of the tag7 gene expressed in metastatic mouse tumors																
JOURNAL	Genetika 32 (5), 621-628 (1996)																
EDITION	MEDLINE																
PUBLISHED	9755036																
REFERENCE	2 (bases 1 to 678)																
AUTHORS	Kustikova, O.S.																
TITLE	Direct Submission																
JOURNAL	Submitted (18-APR-1995) O.S. Kustikova, Inst of Gene Biology RUSSIA																
EDITION	B334, Vavilov Street 34/5, Moscow, RUSSIA																
REFERENCE	3 (bases 1 to 678)																
AUTHORS	Kiselev, S.L., Kustikova, O.S., Korobko, E.V., Prokhortchouk, E.B., Kabishev, A.A., Lukianidin, E.M. and Georgiev, G.P.																
TITLE	Molecular cloning and characterization of the mouse tag7 gene encoding a novel cytokine																
JOURNAL	J. Biol. Chem. 273 (29), 18633-18639 (1998)																
EDITION	MEDLINE																
PUBLISHED	9660837																
COMMENT	On Jul 6, 1995 this sequence version replaced gi:785010.																
FEATURES	Location/Qualifiers																
SOURCE	1. 678 Location/Qualifiers																
ORGANISM	/organism="Mus musculus"																
MATERIAL	/strain="A/Sn"																
REFERENCE	/db_xref="taxon:10090"																
AUTHORS	/tissue_type="adenocarcinoma"																
TITLE	/dev_stage="adult"																
JOURNAL	/pubmed																
EDITION	CDS																
REFERENCE	1. 678 CDS																
AUTHORS	/protein_id="CAA60133..1"																
TITLE	/cell_line="3D c13"																
JOURNAL	/cell_type="myeloblast"																
FEATURES	/codon_start=1																
SOURCE	/product="TAG7-like protein"																
ORGANISM	/protein_id="AAF60325..1"																
MATERIAL	/db_xref="GI:6273361"																
REFERENCE	/db_xref="SPREMBL:Q62185"																
AUTHORS	/translation="MLFACALLAAGLGLATSCSFIVPRSEWRALPSECSSRLGHGPVRYVISHTAGSFCNSDSCQARVQHKNELGMCDVAYFLIGEDGHYEGRGNWKGDIRGPINPMSTICGFMNFDRVPAKRALRNLLECGVSRLERSNEYKHDRVQSTLSPDQLQVOSWHEYR"																
TITLE	/db_xref="DHPGPIWNMSIGTFEMGNFDRVPAKRALRNLLECGVSRLERSNEYKHDRVQSTLSPDQLQVOSWHEYR"																
JOURNAL	/db_xref="DHPGPIWNMSIGTFEMGNFDRVPAKRALRNLLECGVSRLERSNEYKHDRVQSTLSPDQLQVOSWHEYR"																
FEATURES	/codon_start=1																
SOURCE	/protein_id="CAAA60133..1"																
ORGANISM	/cell_line="3D c13"																
MATERIAL	/cell_type="myeloblast"																
REFERENCE	/codon_start=1																
AUTHORS	/product="TAG7-like protein"																
TITLE	/protein_id="AAF60325..1"																
JOURNAL	/db_xref="GI:6273361"																
FEATURES	/codon_start=1																
SOURCE	/protein_id="SPREMBL:Q62185"																
ORGANISM	/translation="MLFACALLAAGLGLATSCSFIVPRSEWRALPSECSSRLGHGPVRYVISHTAGSFCNSDSCQARVQHKNELGMCDVAYFLIGEDGHYEGRGNWKGDIRGPINPMSTICGFMNFDRVPAKRALRNLLECGVSRLERSNEYKHDRVQSTLSPDQLQVOSWHEYR"																
MATERIAL	/db_xref="DHPGPIWNMSIGTFEMGNFDRVPAKRALRNLLECGVSRLERSNEYKHDRVQSTLSPDQLQVOSWHEYR"																
REFERENCE	/codon_start=1																
AUTHORS	/product="TAG7-like protein"																
TITLE	/protein_id="AAF60325..1"																
JOURNAL	/db_xref="GI:6273361"																
FEATURES	/codon_start=1																
SOURCE	/protein_id="SPREMBL:Q62185"																
ORGANISM	/translation="MLFACALLAAGLGLATSCSFIVPRSEWRALPSECSSRLGHGPVRYVISHTAGSFCNSDSCQARVQHKNELGMCDVAYFLIGEDGHYEGRGNWKGDIRGPINPMSTICGFMNFDRVPAKRALRNLLECGVSRLERSNEYKHDRVQSTLSPDQLQVOSWHEYR"																
MATERIAL	/db_xref="DHPGPIWNMSIGTFEMGNFDRVPAKRALRNLLECGVSRLERSNEYKHDRVQSTLSPDQLQVOSWHEYR"																
REFERENCE	/codon_start=1																
AUTHORS	/product="TAG7-like protein"																
TITLE	/protein_id="AAF60325..1"																
JOURNAL	/db_xref="GI:6273361"																
FEATURES	/codon_start=1																
SOURCE	/protein_id="SPREMBL:Q62185"																
ORGANISM	/translation="MLFACALLAAGLGLATSCSFIVPRSEWRALPSECSSRLGHGPVRYVISHTAGSFCNSDSCQARVQHKNELGMCDVAYFLIGEDGHYEGRGNWKGDIRGPINPMSTICGFMNFDRVPAKRALRNLLECGVSRLERSNEYKHDRVQSTLSPDQLQVOSWHEYR"																
MATERIAL	/db_xref="DHPGPIWNMSIGTFEMGNFDRVPAKRALRNLLECGVSRLERSNEYKHDRVQSTLSPDQLQVOSWHEYR"																
REFERENCE	/codon_start=1																
AUTHORS	/product="TAG7-like protein"																
TITLE	/protein_id="AAF60325..1"																
JOURNAL	/db_xref="GI:6273361"																
FEATURES	/codon_start=1																
SOURCE	/protein_id="SPREMBL:Q62185"																
ORGANISM	/translation="MLFACALLAAGLGLATSCSFIVPRSEWRALPSECSSRLGHGPVRYVISHTAGSFCNSDSCQARVQHKNELGMCDVAYFLIGEDGHYEGRGNWKGDIRGPINPMSTICGFMNFDRVPAKRALRNLLECGVSRLERSNEYKHDRVQSTLSPDQLQVOSWHEYR"																
MATERIAL	/db_xref="DHPGPIWNMSIGTFEMGNFDRVPAKRALRNLLECGVSRLERSNEYKHDRVQSTLSPDQLQVOSWHEYR"																
REFERENCE	/codon_start=1																
AUTHORS	/product="TAG7-like protein"																
TITLE	/protein_id="AAF60325..1"																
JOURNAL	/db_xref="GI:6273361"																
FEATURES	/codon_start=1																
SOURCE	/protein_id="SPREMBL:Q62185"																
ORGANISM	/translation="MLFACALLAAGLGLATSCSFIVPRSEWRALPSECSSRLGHGPVRYVISHTAGSFCNSDSCQARVQHKNELGMCDVAYFLIGEDGHYEGRGNWKGDIRGPINPMSTICGFMNFDRVPAKRALRNLLECGVSRLERSNEYKHDRVQSTLSPDQLQVOSWHEYR"																
MATERIAL	/db_xref="DHPGPIWNMSIGTFEMGNFDRVPAKRALRNLLECGVSRLERSNEYKHDRVQSTLSPDQLQVOSWHEYR"																
REFERENCE	/codon_start=1																
AUTHORS	/product="TAG7-like protein"																
TITLE	/protein_id="AAF60325..1"																
JOURNAL	/db_xref="GI:6273361"																

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on : November 4, 2002, 18:59:21 ; Search time 2209 Seconds
 (without alignments)
 117,305 Million cell updates/sec

Title: US-09-462-625-2_COPY_145_160
 Perfect score: 86
 Sequence: 1 RGFLRSNEYVKGHRDV 16

Scoring table: BLOSUM62
 Xgapext 10.0 , Xgapext 0.5 .
 Ygapext 10.0 , Ygapext 0.5 .
 Fgapext 6.0 , Fgapext 7.0 .
 Delext 6.0 , Delext 7.0 .

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000
 Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Command line parameters:
 MODEL=frame+p2n.model -DEV=exp1
 -Q=/cn2-1/USPTO/spool/US09462625/runat_04112002_110248_23136/app_query.fasta_1.199
 -DB=EST -QFMT=fastaP -SUFFIX=_first -MINMATCH=0.1 -LOOPEXT=0
 -UNITS=bits -BITS START=1 -END=1
 -DOCALIGN=200 -THR SCORE=pct -THR MAX=1000 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
 -OUTEXT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 -USER=US09462625 -CGN_1.1_1716_trunat_04112002_110248_23136 -NCPU=6 -ICPU=3
 -NO_XLIPXY -NO_MMAPP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
 -WARN TIMEOUT=30 -THREADS=1 -XGAPPOP=10 -NGAPEXT=0.5 -FGAPEXT=7
 -YGAP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST;*

1: em_estba:*

2: em_estthm:*

3: em_estin:*

4: em_esimu:*

5: em_estcov:*

6: em_estpl:*

7: em_htco:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htcc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_liv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fur:*

23: em_gss_man:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rnd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	86	100.0	274	13 BT159353	BT159353 602919837
2	86	100.0	277	9 AV085455	AV085455
3	86	100.0	281	9 AA691322	AA691322 vs14901..x
4	86	100.0	305	9 AA910493	AA910493 vo45e06..x
5	86	100.0	316	9 AA910504	AA910504 vo45e09..x
C	6	86	100.0	322	9 AI849253 UT-M-AI-1
					AA616255 vo50e07..x
					AA689633 vs09g08..x
					AA724993 vs17f04..x
					AA734805 vp38g01..x
					AA889693 vs03d09..x
					AI895767 vs03d09..x
					AA897240 vo28c1..x
					BG693242 602828323
					AIF07116 vj8209..x
					B1556032 60237772
					BG174272 602334571
					AA238564 my35e0..x
					BG871384 602919056
					B1154844 602802826
					BE119698 ug5c1..x
					BF302505 60203134
					BI149595 602848402
					BI155774 602804343
					BT165038 60297736
					AK008335 Mus muscu
					BQ951956 AGENCOURT
					B6244455 602356590
					BI453419 603170810
					BF163190 601771917
					BT1650815 602861906
					AV092014 AV092014
					BG975104 602843369
					AA38752 my35e09..x
					AV082572 AV082572
					BL291182 UI-R-DK0-
					BP548605 UI-R-A0-a
					BA875213 UI-R-E0-c
					AI153056 ud53a07..x
					AA228200 SWMMCA112
					BF076851 226604 MA
					BQ564677 9121b09..y
					BB654595 UT-M-AU-1
					AV065626 AV065626
					AV062378 AV062378

ALIGNMENTS

RESULT 1	BT159353	274 bp	mRNA	linear	EST 05-JUL-2001
LOCUS	602919837E1 NCI_CGAP_Man3	Mus musculus cDNA clone	IMAGE:5.060072 5'	mRNA sequence.	
DEFINITION					
ACCESSION	BT159353.1	GI:14619354			
VERSION					
KEYWORDS					
SOURCE				house mouse.	
ORGANISM				Mus musculus	
				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
				Mammalia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE	1 (bases 1 to 274)			NIH-NGC http://mgc.ncbi.nlm.gov/	
AUTHORS				National Institutes of Health, Mammalian Gene Collection (MGSC)	
TITLE					

Fax: 81-298-36-9098
Email: genome-res@tc.riken.go.jp
The thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA (Proc. Natl. Acad. Sci. U.S.A. 95(2): 520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (<http://genome rtc.riken.go.jp>) for further details.

Pax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LILN ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Mgi:619496
 Putative full length read
 vector to vector length is
 Seq primer: -28m13 rev2 ET from Amersham.
 High quality sequence stop: 203.

FEATURES

source	Location/Qualifiers
	1. .281 <code>/organism="Mus musculus"</code> <code>/strain="FVB/N"</code> <code>/db_xref="taxon:1090"</code> <code>/clone="IMAGE:1130224"</code> <code>/clone_lib="Barstead mouse irradiated colon MPLRB7"</code> <code>/dev_stage="8 weeks"</code> <code>/lab_host="DH10B"</code>
	<code>/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained from 8 week old mouse. Colon was harvested 72 hours after irradiation with 1400 Gys. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TCCTTACGAATCTGAACTGGCGCCCTTGTTCCTTTTTTTTTT T 3'-]; double stranded cDNA was ligated to Eco RI adaptors [AATTGGATCTTGAGCTGGCGCCGCTTGTTCCTTTTTTTT T 3'-]; double stranded cDNA was digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed by Bob Barstead."</code>

BASE COUNT
 ORIGIN

69 a	87 c	67 g	58 t
------	------	------	------

Alignment Scores:
 Pred. No.: 3.89e-07
 Score: 86.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 9
 DB:

US-09-462-625-2_COPY_145_160 (1-16) x AA930493 (1-305)
 Qy 1 ArgGlyPheLeuArgSerAsnTyrGluValLysGlyHisArgAspVal 16
 Db 96 CGGGCCTCTGAGATCCAACTATGAAGTCAAAGCACGGGATGTG 143

RESULT 4
 AA930493
 LOCUS vo45c06_r1
 DEFINITION Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA clone IMAGE:1052842 5' similar to gb:X86374 M.musculus mRNA for TAG7 protein (MOUSE); mRNA sequence.

ACCESSION AA930493
 VERSION 1 GI:3079983
 SOURCE EST
 ORGANISM Mus musculus

Mus musculus; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus. 1 (bases 1 to 305)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Scheibenber,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Watson,R.

TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra,M/Mouse EST Project
 WashU HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu

TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra,M/Mouse EST Project
 WashU HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LILN ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:584418
 putative full length read
 vector to vector length is 306
 Seq primer: -28m13 rev2 ET from Amersham.
FEATURES

source	Location/Qualifiers
	1. .305 <code>/organism="Mus musculus"</code> <code>/strain="FVB/N"</code> <code>/db_xref="taxon:10090"</code> <code>/clone="IMAGE:1052842"</code> <code>/clone_lib="Barstead mouse irradiated colon MPLRB7"</code> <code>/dev_stage="8 weeks"</code> <code>/lab_host="DH10B"</code>
	<code>/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained from 8 week old mouse. Colon was harvested 72 hours after irradiation with 1400 Gys. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TCCTTACGAATCTGAACTGGCGCCCTTGTTCCTTTTTTTT T 3'-]; double stranded cDNA was ligated to Eco RI adaptors [AATTGGATCTTGAGATCCAACTATGAAGTCAAAGCACGGGATGTG 143</code>

BASE COUNT
 ORIGIN

74 a	97 c	69 g	65 t
------	------	------	------

Alignment Scores:
 Pred. No.: 4.42e-07
 Score: 86.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 9
 DB: 100.00%
 US-09-462-625-2_COPY_145_160 (1-16) x AA930493 (1-305)
 Qy 1 ArgGlyPheLeuArgSerAsnTyrGluValLysGlyHisArgAspVal 16
 Db 96 CGGGCCTCTGAGATCCAACTATGAAGTCAAAGCACGGGATGTG 143

RESULT 5
 AA930504
 LOCUS AA930504
 DEFINITION AA930504_r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA clone IMAGE:1052849 5' similar to gb:X86374 M.musculus mRNA for TAG7 protein (MOUSE); mRNA sequence.

ACCESSION AA930504
 VERSION 1 GI:3079994
 KEYWORDS EST
 SOURCE house mouse
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 316)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Scheibenber,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Watson,R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra,M/Mouse EST Project
 WashU HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LILN ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:584425

/lab_host="DH10B"
 /note="Vector: pT7TD-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained from 8 week old mouse. Colon was harvested 72 hours after irradiation with 1400 Gys. 1st strand cDNA was primed with a Not I - oligo(dt) primer [5'-TGTTAGAATCTGGAACTGGGCCCTTGTGTTTTTTTTTTT]
 T 3'; double stranded cDNA was ligated to Eco RI adaptors [AATTCGATCTTG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library constructed by Bob Barstead."

BASE COUNT 107 a 153 c 149 g 119 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.04e-06 Length: 528
 Score: 86.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: DB: US-09-462-625-2_COPY_145_160 (1-16) x AA734805 (1-534)

QY 1 Arg Gly Phe Ieu Arg Ser Asn Tyr Glu Val Lys Gly His Arg Asp Val 16
 Db 470 CGGGCCTTCCTGAGATCCAACATGAAGTCAAAGGAAACCGGAATCTG 517

RESULT 10
 AA734805 LOCUS AA734805 534 bp mRNA Linear EST 07-JAN-1998
 DEFINITION vp8b01.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus CDNA clone IMAGE:1078921 5' similar to gb:x86374 M.musculus mRNA for TAG7 protein (MOUSE);, mRNA sequence.

AA734805.1 EST.
 AA734805.1 GI:2756472
 VERSION
 SOURCE
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.,
 REFERENCE 1 (bases 1 to 547)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Scheibenbogen, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE The WashU-HMMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 Washington University School of MedicineP
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

MG:593217 TITLE
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 Washington University School of MedicineP
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES Source
 1. .534 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:1078921"
 /clone_id="Barstead mouse irradiated colon MPLRB7"
 /dev_stage="8 weeks"
 /lab_host="DH10B"
 /note="Vector: pT7TD-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained from 8 week old mouse. Colon was harvested 72 hours after irradiation with 1400 Gys. 1st strand cDNA was primed with a Not I - oligo(dt) primer [5'-TGTTAGAATCTGGAACTGGGCCCTTGTGTTTTTTTTT]
 T 3'; double stranded cDNA was ligated to Eco RI

irradiation with 1400 Gys. 1st strand cDNA was primed with a Not I - oligo(dt) primer [5'-TGTTAGAATCTGGAACTGGGCCCTTGTGTTTTTTTTT]
 T 3'; double stranded cDNA was ligated to Eco RI adaptors [AATTCGATCTTG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library constructed by Bob Barstead."

BASE COUNT 1130 a 156 c 137 g 111 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.06e-06 Length: 534
 Score: 86.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: DB: US-09-462-625-2_COPY_145_160 (1-16) x AA734805 (1-534)

QY 1 Arg Gly Phe Ieu Arg Ser Asn Tyr Glu Val Lys Gly His Arg Asp Val 16
 Db 326 CGGGCTCTGAGATCAACTATGAAGTCAAAGGAAACCGGATGTC 373

RESULT 11
 AA689693 LOCUS AA689693 547 bp mRNA linear EST 16-DEC-1997
 DEFINITION vs03d09.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus CDNA clone IMAGE:1137137 5' similar to gb:x86374 M.musculus mRNA for TAG7 protein (MOUSE);, mRNA sequence.

AA689693.1 GI:2690629
 ACCESSION AA689693.1
 VERSION
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.,
 REFERENCE 1 (bases 1 to 547)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Scheibenbogen, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE The WashU-HMMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 Washington University School of MedicineP
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

MG:618409 SEQ PRIMER Seq primer: -28m13 rev2 ET from Amersham High quality sequence stop: 141.

FEATURES Source
 Location/Qualifiers
 1. .534 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:1078921"
 /clone_id="Barstead mouse irradiated colon MPLRB7"
 /dev_stage="8 weeks"
 /lab_host="DH10B"
 /note="Vector: pT7TD-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained from 8 week old mouse. Colon was harvested 72 hours after irradiation with 1400 Gys. 1st strand cDNA was primed with a Not I - oligo(dt) primer [5'-TGTTAGAATCTGGAACTGGGCCCTTGTGTTTTTTT]
 T 3'; double stranded cDNA was ligated to Eco RI

Alignment Scores:								
Pred. No. :	1.1e-06	Length:	548					
Score:	86.00	Matches:	16					
Percent Similarity:	100.00%	Conservative:	0					
Best Local Similarity:	100.00%	Mismatches:	0					
Query Match:	100.00%	Indels:	0					
DB:	9	Gaps:	0					
	us-09-462-625-2_COPY_145_160 (1-16) x AA597240 (1-548)							
Qy	1 Arg Gly Phe Leu Arg Ser Asn Tyr Glu Val Lys Gly His Arg Asp Val 16							
Db	342 CGGGGCTCTTGTAGATCCAACTATGAAGTCAGACCGGGATGTG 389							
RESULT 14								
BG963242	BG963242 571 bp mRNA linear EST 12-JUN-2001							
DEFINITION	60282823F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4983340 5'							
mRNA sequence.								
ACCESSION	BC963242							
VERSION	EST.							
KEYWORDS								
SOURCE	Mus musculus							
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.							
COMMENT	(bases 1 to 575)							
REFERENCE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.							
AUTHORS	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),							
TITLE	Tumor Gene Index							
JOURNAL	Unpublished (1997)							
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabps-r@mail.nih.gov							
REFERENCE	This clone is available royalty-free through LILN; contact the IMAGE Consortium (info@image.llnl.gov) for further information.							
AUTHORS	MGI:540481							
TITLE	MGIC http://mgc.nci.nih.gov/							
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)							
COMMENT	Unpublished (1999)							
CONTACT	Contact: Robert Strausberg, Ph.D.							
REFERENCE	Email: cgabps-r@mail.nih.gov							
TITLE	Tissue Procurement: Jeffrey E. Green, M.D.							
JOURNAL	cDNA Library Preparation: Life Technologies, Inc.							
COMMENT	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILN)							
CONTACT	DNA Sequencing by: Incyte Genomics, Inc.							
REFERENCE	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILN at: http://image.llnl.gov							
TITLE	Plate: LILN01987 row: m column: 05							
JOURNAL	High quality sequence stop: 312.							
COMMENT	Location/Qualifiers							
FEATURES	1. 571							
SOURCE	/organism="Mus musculus"							
COMMENT	/strain="FVB/N"							
REFERENCE	/db_xref="taxon:10090"							
TITLE	/clone="IMAGE:4983340"							
JOURNAL	/clone_lib="NCI_CGAP_Co24"							
COMMENT	/lab_host="DH10B (T1 phage-resistant)"							
CONTACT	/note="Organ: colon; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: Sall; Cloned unidirectionally; Primer: Oligo dR.							
REFERENCE	Average insert size 1.6 kb. Constructed by Life							
TITLE	Technologies. Note: this is a NCI_CGAP Library."							
JOURNAL	BASE COUNT 199 a 203 c 111 g 58 t							
COMMENT	ORIGIN							
FEATURES	1.17e-06	Length:	571					
SOURCE	86.00	Matches:	16					
COMMENT	Percent Similarity: 100.00%	Conservative:	0					
REFERENCE	Best Local Similarity: 100.00%	Mismatches:	0					
TITLE	Query Match: 100.00%	Indels:	0					
JOURNAL	DB: 13	Gaps:	0					
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
SOURCE								
COMMENT								
REFERENCE								
TITLE								
JOURNAL</td								